Influence of Cancer Associated Microbiome on Volatile Organic Compound Production in Oesophago-gastric Adenocarcinoma

MINA SAMY EDWARD ADAM

Department of Surgery and Cancer, Imperial College London

St Mary's Hospital, London, UK

2019

A thesis submitted to Imperial College London for the degree of

Doctorate of Philosophy

Supervisors

Professor George B Hanna PhD FRCS

Department of Surgery and Cancer, Imperial College London,

St Mary's Hospital, London, UK

Professor Julian Marchesi PhD

Department of Surgery and Cancer, Imperial College London,

St Mary's Hospital, London, UK

THESIS DECLARATION

I hereby declare that I am the sole author of this thesis and all work contained within it is my own. All individuals who collaborated on any sections of the work are appropriately credit.

Signed: Mina Adam

Date: 21/08/2019

DEDICATION

This thesis is dedicated to my family and in particular my parents. Words of gratitude and appreciation are not enough to thank them for being incredibly supportive throughout, and have made so many sacrifices for my training and career. Thank you for your patience, love and endless encouragement.

ACKNOWLEDGEMENTS

Throughout the period of undertaking this research project I have received support from a large number of individuals who have enabled me to complete this work. Firstly, I am very fortunate to have had a supervisor as outstanding as Professor George B Hanna whom I am indebted for his tremendous support. His patience, motivation and constant guidance has been fundamental to the successful completion of this work. As an international student, not familiar with the UK system, Professor Hanna has always provided me with encouragement, patience and guidance to the course of my studies. This is a testament of his unique character. I am also very grateful to my other supervisors, Professor Julian Marchesi and Mr Sacheen Kumar, both of whom provided me with invaluable advice, encouragement and expert insights during my research period.

My greatest thanks must go to Mr Piers Boshier, as he has been a hugely supportive mentor and brilliant teacher, and I am extremely grateful for all his support throughout. No matter whatever obstacles there were to progress, Piers has always given me huge amounts of time towards making this project successful. I am also grateful to my Clinical Research Fellow colleagues with particular thanks to Mr Pranav Patel especially for his help with the tissue bank, Mr Matyas Ferhevari, Mr Tom Wiggins and Mr Stefan Antonowicz.

Other thanks for friendship, assistance and solidarity in the lab are owed to Dr Ilaria Belluomo, Dr SungTong Chin, Dr Andrea Romano, Dr Oscar Ayron and Dr Julie McDonald. This work would not have been possible without their fantastic support. I would like to thank Mr. Krishna Moorthy, Mr Christopher Peters, Dr. Alison Knaggs and the St. Mary's

Hospital Endoscopy Staff for allowing the recruitment of patients under their care.

Finally, I was fortunate to collaborate with Ingenza, who supported bacterial culture studies and development of the oral stimulant drink. Dr Ian Fotheringham, Dr Claire Wilson, Dr Fraser Brown, Dr Alison Arnold, Dr Micheal Herrera, all of them are inspiring scientists and I am grateful for their immense support.

Abstract

Oesophago-gastric cancer is a significant health problem with poor prognosis in Western countries. This is due to a paucity of alarm symptoms in early stages of the disease resulting in late clinical presentation and associated delays in initiation of treatment. The development of non-invasive breath tests using exhaled Volatile Organic compounds (VOCs) to determine oesophago-gastric cancer risk would help facilitate earlier diagnosis and potentially improve patient survival. Whilst many of the biochemical pathways relating to the origin of these VOCs within humans are as yet unknown, it is postulated that that specific VOCs are produced directly by cancer tissues. Contributions from other endogenous sources including the intestinal microbiome and healthy tissues within the intestinal tract as well as other organ systems. The aim of this thesis was to understand the interaction between the upper gastrointestinal microbiome and VOC production in patients with oesophago-gastric cancer and to explore how this onco-microbial axis can be exploited to augment VOC production.

The production of cancer associated VOCs (fatty acids and phenol) were investigated by analysing the *ex vivo* headspace above un-derivatised tissue samples as well as *in vivo* mixed breath, isolated bronchial breath and gastric endoluminal air. Increased concentrations of these VOCs were detected in the headspace of cancer tissue samples as well as isolated endoluminal air adjacent to tumours. Findings therefore implicate that the tumour and its local environment are the likely source of upregulated VOCs in oesophago-gastric cancer. The relative contribution of the tumour associated microbiome remains unknown.

16S RNA sequencing analysis for 185 oesophago-gastric tissue samples from cancer and control subjects were performed in order to assess the microbial diversity. Results revealed higher abundance of Firmicutes (e.g. Streptococcus salivarius, Escherichia coli and Streptococcus anginosus) in oesophago-gastric cancer samples compared to controls. The headspace of *in vitro* and patient derived (*ex vivo*) cultures of specific targeted bacteria was subsequently found to contain similar VOCs as those previously detected in oesophago-gastric cancer.

To increase the sensitivity of breath testing, further work was performed to augment the diagnostic response using simple metabolic substrates (sugars, proteins, lipids). When added to *in vitro* cultures of cancer-associated bacteria, these nutrients resulted in upregulated VOC production. Oesophago-gastric cancer patients who were given the same substrates orally were found to have a transient rise in the same VOCs that was greater than observed in healthy controls.

This thesis provides new insight into the biological origin of VOC production in oesophago-gastric cancer. Experiments linking the cancer-associated microbiome, exogenous substrates to upregulated VOC production in cancer patients offers the potential for a future augmented breath test for this disease. The augmented breath test is expected to increase earlier cancer detection leading to improvement in overall survival.

PRIZES, PRESENTATIONS AND PUBLICATIONS

Prizes and grants awarded for scientific research

1. European Society of Surgery Annual meeting Poster of distinction prize (2017)
2. Mass Spectrometry Applications to the clinical lab young investigator Fellowship (MSACL Jan 2018)
3. Mass Spectrometry Applications to the clinical lab young investigator Fellowship (MSACL Sep 2018)
4. Mass Spectrometry Applications to the clinical lab young investigator Fellowship (MSACL 2017)
5. London Surgical Symposium Poster of Distinction (2017)
6. European Society of Surgery Poster of Distinction (2017)

Publications in peer reviewed journals

1. Mina E Adam, Matyas Fehervari, Piers R Boshier *et al*. Mass spectrometry analysis of mixed breath, isolated bronchial breath and gastric endoluminal volatile fatty acids in oesophago-gastric cancer. Analytical chemistry. 2019 ;91(5):3740-3746.

Publications in Conference extracts

1. In vitro volatile compound bacterial profiling using GC-MS and their Potential Role in the Diagnosis of Oesophago-Gastric Cancer. Oral podium presentation at the International Mass Spectrometry Application to the clinical Lab conference, MSACL Palm Springs US (2018).
2. Proton Transfer Reaction time-of-flight mass spectrometry analysis of volatile organic compounds in human gastric cancer Tissue. Oral podium presentation at the International Mass Spectrometry Application to the clinical Lab conference, MSACL Salzburg (2017).

Presentations to learned societies

1. In vitro volatile compound bacterial profiling using GC-MS and their Potential Role in the Diagnosis of Oesophago-gastric Cancer. Oral podium presentation at the International Mass Spectrometry Application to the clinical Lab conference, MSACL Palm Springs US (2018).
2. Characterization of volatiles and identification of gastro-oesophageal cancer specific

endoscopic tissue headspace compounds. Oral podium presentation at the European Association for Endoscopic Surgery, London (2018).

3. Culturing gastro-oesophageal cancer tissue and their potential diagnostic and therapeutic significance. Oral podium presentation at the European Society for Surgical Research, Madrid (2018).
4. Proton Transfer Reaction time-of-flight mass spectrometry analysis of volatile organic compounds in human gastric cancer Tissue. Oral podium presentation at the International Mass Spectrometry Application to the clinical Lab conference, MSACL Salzburg (2017).
5. Assessment of microbial diversity and the effects on oesophago-gastric adenocarcinoma. Oral podium presentation at the European Society of Surgery, Annual Meeting, ESS Poland (2017).
6. Headspace analysis of volatile organic compounds in human gastro-oesophageal adenocarcinoma. Poster of distinction at the European Society of Surgery, Annual meeting, ESS Poland (2017).
7. The influence of the microbiome and effects on volatile organic compound production in oesophago-gastric adenocarcinoma. Poster of distinction at the London Surgical Symposium (2017).
8. Assessment of microbial diversity and its effects on volatile compound production in oesophago-gastric adenocarcinoma. Poster presentation at Imperial College Graduate School (2016).

CONTENTS

1 LIST OF TABLES

2 LIST OF FIGURES

3 LIST OF ABBREVIATIONS

GPORD – Gastro-Oesophageal Reflux Disease
GC-MS Gas-Chromatography Mass Spectrometry
GI – gastrointestinal
H.pylori – Helicobacter pylori
HCl – Hydrochloric acid
IQR – Inter-quartile range
LC-MS – Liquid chromatography mass spectrometry
LOD – Limit of detection
LOQ – Limit of quantification
m/z – Mass-to-charge ratio
MWU – Mann-Whitney U
MALDI-TOF-MS–Matrix assisted laser desorption/ionization-time of flight-mass spectrometry
NUD– Non ulcer dyspepsia
OG Cancer– Oesophago-Gastric Cancer
OGD – Oesophago-Gastro-Duodenoscopy
PBS– Phosphate Buffer Solution
ppbv – parts-per-billion by volume
pptv – parts-per-trillion by volume
PTR-MS– Proton Transfer Reaction Mass Spectrometry
PUD–Peptic ulcer disease
qPCR – Quantitative polymerase chain reaction
RT– Retention Time
RNA – ribonucleic acid
ROC – Receiver operating characteristic
SIFT-MS– Selected Ion Flow Tube Mass Spectrometry
SPME – Solid phase microextraction
S/N – Signal to noise ratio
T-RFLP– Terminal resection fragment length polymorphism
TD– Thermal Desorption
VOC– Volatile Organic Compound

4 INTRODUCTION

4.1 Oesophago-gastric cancer

Oesophago-gastric (OG) cancer is currently the sixth leading cause of cancer related death worldwide. In 2008, there were 482,000 new cases of oesophago-gastric cancer diagnosed and 407,000 oesophago-gastric cancer deaths recorded globally[1].

Oesophageal cancer is divided into two major histological subtypes. Squamous cell carcinoma is more common within Central and South-East Asia whereas oesophageal adenocarcinoma is more prevalent in Western Countries. In 2012, there were 398,000 cases of squamous cell carcinoma and 52,000 cases of oesophageal adenocarcinoma worldwide. The overall global incidence of oesophageal cancer is 0.7 per 100,000 (1.1 per 100,000 in men and 0.3 per 100,000 in women)[2]. The incidence of gastro-oesophageal junctional adenocarcinoma has significantly increased in Europe and the USA in the past thirty years with both reflux oesophagitis and obesity being primary risk factors contributing to this increase[3,4,5,6,7,8]. The poor prognosis of OG adenocarcinoma is defined by it age-standardised 5-year relative survival rate of 13% in the UK[9].

The Incidence of gastric cancer is mostly highest in eastern Asia (particularly China, Japan Korea, and), Mongolia) South America and Central and Eastern Europe, whilst being lowest in most parts of Africa and Northern America[10] .The predominant gastric cancer subtype is adenocarcinoma[11]. The incidence of gastric cancer is known for its large differences that exist geographically with an incidence of up to 3.4 per 100 000 among females of North America, compared to 26.9 per 100 000 among males in Asia[12]. Chronic infection with

Helicobacter Pylori is a known risk factor for stomach cancer with 89% of non-cardia gastric cancer cases associated with infection with this bacterium[13].

In the UK, upper gastrointestinal symptoms account for at least 3% of consultations in primary care[14]. Oesophago-gastric cancers often present to clinicians with a spectra of non-specific symptoms & signs including nausea, vomiting, abdominal pain, dysphagia, dyspepsia, odynophagia and/or weight loss. The variability of the early signs of oesophago-gastric cancer and lack of 'alarm' symptoms until the disease is at a more advanced stage result in only one-third of patients being suitable for treatment with curative intent at diagnosis[15]. However, definitive treatment of early (stage I) OG adenocarcinoma (either endoscopically or surgically) has a reported 5-year survival rate of 83-96%.[16] In Japan, the national screening programme helps facilitates early detection of gastric cancer disease with better oncological outcomes[17]. At present there are no national screening programmes for oesophago-gastric cancer in the majority of countries, largely because the incidence is relatively low and that the gold-standard diagnostic test (endoscopy) is expensive, and due to the invasive nature of this test there are potential risks to patients.

There remains therefore an unmet clinical need to develop novel diagnostic techniques as well as increasing our understanding of the underlying carcinogenic mechanisms in ooesophago-gastric cancer to improve early detection and long-term survival. Analysis of Volatile Organic Compounds (VOCs) within exhaled breath of oesophago-gastric cancer patients has been proposed as one potential method of risk stratification prior to endoscopy.

4.2 **Volatile Organic Compounds (VOCs)**

Volatile Organic Compounds are carbon-based chemicals that are predominantly gaseous at ambient temperature. VOCs emitted from the human body have been of interest to researchers for several decades. In 1971, Pauling identified more than 250 and 280 VOCs within exhaled breath and urine, respectively in healthy human subjects[18]. The identification, characterisation and accurate quantification of VOCs however remains challenging. Various direct injection mass spectrometry techniques are available for VOC analysis including Selected-Ion Flow-Tube Mass Spectrometry (SIFT-MS) and Proton-Transfer Reaction Mass Spectrometry (PTR-MS). However, Gas Chromatography-Mass Spectrometry (GC-MS) is still the most widely employed platform despite its limitations for identifying VOCs with low molecular weights.

4.2.1 Results of Preliminary VOC Research in oesophago-gastric cancer

A number of previous studies have sought to define a unique VOC profile associated with oesophago-gastric cancer[19-24]. In a cohort of 210 patients referred for upper gastrointestinal endoscopy (81 cancer and 129 controls), VOCs within exhaled breath were evaluated using Selected Ion Flow Tube Mass Spectrometry (SIFT-MS) in oesophago-gastric cancer[21]. Twelve VOCs were present at significantly higher concentrations in both cancer groups compared to non-cancer controls. These results were not explained by potential confounding factors that were examined separately. A diagnostic prediction model to discriminate oesophago-gastric adenocarcinoma (N = 81) from non-cancer controls (including Barrett's metaplasia, benign group & normal UGI tract N = 129) was constructed. The eight most significant predictors for adenocarcinoma identified from stepwise logistic regression were: decanal, nonanal,

heptanal (i.e., aldehydes), phenol, ethyl phenol, methyl phenol (i.e., phenols), hexanoic acid and butyric acid (i.e., fatty acids). After being subjected to Monte Carlo simulation, the AUC curve for the model subset was 0.92 ± 0.01 with a sensitivity of 89.3% (95% CI, 77.0 – 95.7) and specificity of 83.7% (95% CI, 74.5 – 90.9). The AUC for the validation subset was 0.87 ± 0.03, with a sensitivity of 86.7% (95% CI, 71.6 – 97.6) and specificity of 81.2% (95% CI, 64.0 – 90.0). The model subset, on average, correctly classified 120 out of 139 patients (diagnostic accuracy 86.0% ± 0.02) with 6 false negatives. The validation subset correctly classified 59 out of 71 patients (diagnostic accuracy 83.3% ± 0.04), with 3 false negatives.

These findings were broadly replicated in a larger multicentre validation study. Based on five principal VOCs (butyric acid, hexanoic acid, pentanoic acid, Butanal and decanal) the model demonstrated a sensitivity of 80%, specificity of 81% and area under the receiver operator characteristic curve 0.85. Collectively these findings highlight the potential for VOC to be used as non-invasive biomarkers of oesophago-gastric cancer[19]. Whilst many of the biochemical pathways relating to the origin of these VOCs within humans are as yet unknown, it is postulated that that specific VOCs are produced directly by cancer tissues. Contributions from other endogenous sources including the intestinal microbiome and healthy tissues within the intestinal tract as well as other organ systems.

4.3 **Overview of Mass Spectrometry**

Mass spectrometry has been the primary analytical techniques used for the identification and quantification of VOCs within exhaled breath. Based on the creation of a unique spectrum of product ions, mass spectrometry can be used to determine the molecular

composition of a range of sample types[25]. Current uses of clinical mass spectrometry include: toxicology testing in body fluids, diagnosis of metabolism deficiencies and for measurement of biomarkers, enzymes or hormone levels[26,27]. Highly sensitive mass spectrometry techniques, such as Proton Transfer Reaction-Time of Flight-Mass spectrometry (PTR-TOF-MS), Gas Chromatography-Mass Spectrometry (GC-MS) and selective ion flow transfer (SIFT-MS) have been developed to measure VOCs and potentially be used to differentiate between the progression of specific diseases [28]. These techniques have enhanced the capability to detect minor differences in compounds and could potentially be employed for use in clinical diagnosis.

4.4 Mass Spectrometry techniques

4.4.1 Gas Chromatography Mass Spectrometry

The basic principle behind GC-MS is the transfer of a sample from its gaseous state from the injector site, which is the heated inlet port, through the column and into the detector via the use of inert carrier gasses[29]. The 'mobile phase' constituents of a mixture of gas can be separated by passage over a solid stationary phase which occurs within the column[25]. Both the separation between the two phases and various interactions between each compound in the stationary phase allows compound identification through the specific retention times[30]. The column is located within an oven having a high stable temperature for the elution process[31]. Compounds interacting more strongly take longer time to reach the detector, thus leading to the production of a plot known as a gas chromatogram[31]. Samples separated by GC-MS are analysed by quadrupole mass detector, attached to the GC column[29].

After a sample is analysed through an instrument, a mass spectrum is generated. A mass spectrum is a graphical representation of the ions over a specified range of *m/z* values[25]. The output is in the form of an *x,y* plot in which the x-axis is the mass-to-charge scale and the y-axis is the intensity scale [25]. If an ion is observed at an *m/z* value, a line is drawn as a signal representing the response of the detector to the product ionic species[29]. The mass spectrum will contain peaks representing both the fragment ions as well as the molecular ion. Interpretation of a mass spectrum identifies, confirms, or determines the quantity of a specific compound[32]. By running standards of tentatively identified compounds using similar settings, peaks at the same retention time are used to confirm the identity of the compound.

Identification and compound separation by GC-MS remains one of the main positives of the technique having few limitations[25]. Although the size of a spectral peak is proportional to the amount of the substance reaching the detector in the GC instrument, it still does not give exact quantification[33]. Development of calibration curves is always needed to allow semi quantification, but developing these presents with technical difficulties. The GC-MS instrument is technically very sensitive and regular maintenance of the instrument is required to rely on the analysis[34].

4.4.2 Proton Transfer Reaction-Mass Spectrometry (PTR-MS)

Proton Transfer Reaction-Time of Flight (ToF)-MS is widely employed in the analysis of Volatile Organic Compounds (VOCs) with the main fields of application being environmental research, food analysis and counter-terrorism[25]. It is a direct injection mass spectrometry technique in which chemical ionization is achieved by using pronated H_2O as the main

reagent ion[35]. The hydronium ions enter a separate compartment from the ion source, called the drift tube, and subsequent ionization is detected by the quadrupole detector[25].

Time-of-flight (ToF) mass analysers can be coupled to PTR-MS instruments to create a high time of resolution and mass resolution [36], which allows for better ion mass quantification [37]. The high time resolution, as a complete mass spectrum, is generated at each ionization event and spectra can be recorded at a frequency ≤ 1Hz [38]. This improves the accuracy of measurement during *in vivo* studies and full spectral scans can be performed without compromising sensitivity, enabling both targeted and untargeted analyses[37].

PTR-MS chemical ionisation allows for a 'softer' ionisation process and low fragmentation, improving mass spectral interpretation[25]. Samples can be analysed directly in their natural physiological state in real time with accurate quantification, without the need for sample preparation pre-analysis[37]. The ionisation process only occurs with molecules with higher affinity than water hence not all compounds can be detected. This issue has been overcome with the use of other carrier ions including O_2^+ and NO^+ with switching between reagent ions possible[35].

4.5 **Overview of the microbiome**

The term microbiome refers to the additional set of genes arising from the diverse and unique array of microbes that have established themselves in a variety of habitats throughout our body. This is not to be confused with the term microbiota, simply referring to their names and quantities. These communities of bacteria, viruses, fungi and yeasts can be found in significant numbers in regions such as the gastrointestinal and reproductive

tracts as well as skin. Overall, the number of microbes in the human body matches the number of native cells at a 1:1 ratio, meaning the microbiome has a significant contribution towards our genetic diversity, harnessing great potential to aid understanding of a number of medical conditions targeted by decades of research[39].

The microbiome is presenting itself as an exciting and rapidly developing field of scientific research. Significant changes in the microbiome have been identified in several diseases including inflammatory bowel disease, obesity, cancers, autism and depression [40]. A feature that all of features have in common is that they have a complex underlying biology and require therapies with multiple targets.

4.5.1 Microbiome in Oesophago-Gastric cancer

The role of the microbiome in tumorigenesis has previously been investigated in Colorectal Cancer (CRC) [41]. Recent data support the idea that some bacterial species could act as "drivers", and the carcinogenic action of these bacteria could either be direct (e.g. production of DNA-damaging compounds) or indirect (e.g. promotion of inflammation) [42]. Limited data is available regarding the upper gastrointestinal microbiome because the stomach was long considered to be a virtually "sterile" environment, primarily due to its highly acidic pH. Gastric content though can have a variable pH range in the presence of upper GI disease. Our group has previously investigated gastric content and observed a median pH of 1.93 for those with normal Upper GI tracts, a median pH of 5.67 for those with benign disease of the Upper GI tract and a median pH of 6.18 in those with oesophago-gastric cancer[23]. Of note, there were no differences in proton pump inhibitor or H_2 receptor antagonist use amongst these three groups. It is well known that *Helicobacter pylori* is an

important risk factor for oesophago-gastric disease through its ability to overcome the acid-mucus barrier and colonise gastric epithelium.[15] However, more recent data demonstrate that other components of the microbiome including; *Proteobacteria*, *Firmicutes*, *Bacteroidetes*, *Fusobacteria*, and *Actinobacteria* are able to survive in the hostile environment around the gastro-oesophageal junction. The effects and influences of these bacterial species with the upper gastrointestinal tract remain largely unknown, particularly in the presence of hypoacidity and disease states[16,43]. Through this research, there is significant potential to identify pathobionts that may be implicated in OG adenocarcinoma and evaluate any potential influence these bacteria may exert on the production of cancer-related Volatile Organic Compounds.

4.5.2 Bacterial production of VOCs

Bacterial strains have a distinct metabolism, which may result in the production of bacterial specific VOCs. Several studies have employed SIFT-MS or gas chromatography (GC) to differentiate between selected bacterial species according to their distinct VOC profile. Real-time headspace analysis was the method of choice to quantify the concentrations of compounds. Findings demonstrated various differences in the VOCs associated with different bacterial strains between specific bacterial strains grown. (Table 1)

It is therefore plausible that dysregulated VOC production in cancer may be in part a consequence of variations in the microbiome.

VOC	Bacteria Identified	Identification methodology	Reference
Butyric acid	E.coli, P.mirabili, B. cepacia, S.pyogenes	SIFT-MS,headspace of monoculture	44
	Actinomyces naeslundii,Clostridium bifermentans,Clostridium perfringens,Clostridium spordenes, Clostridium tertium, Fusobacterium necrphorum, Lactobacillus acidophilus, Peptstreptococcus anaerobius, preptstreptococcus prevotii	Headspace solid phase microextraction combined with gas chromatography	46
Hydrogen sulphide	P. aeruginosa, S. aureus, E. coli, P mirabilis, B. cepacia, E. faecalis	SIFT-MS,headspace of monoculture	44
	S.pneumoniae, E.coli, N.meningitidis	SIFT-MS,headspace of monoculture	45
Pentanol	S.aureus, E.coli, P.mirabili, B.cepaci,S.pyogenes, E.faecalis	SIFT-MS,headspace of monoculture	44
Propanoic acid	Acinetobacter baumannii, Actinomyces naeslundii, Actinomyces naeslundii, Bacteroides pyogenes, Clostridium bifermentans, Clostridium ramosum, Clostridium septicum, Clostridium sporogenes, Clostridum tertium, Escherichia coli, Enterobacter cloacae, Eubacterium lentum, Eubacterium sp., Fusobacterium simiae, Fusobacterium necrophorum, Klebsiella pneumonia, Lactobacillus acidophilus, Nocardia sp.,Peptostreptoccus anaerobius, Preptostreptoccus prevotii	Headpsace solid phase microextraction combined with gas chromatography	46

Table 1:Studies demonstrating the ability of various bacteria to produce specific VOCs identified using mass spectrometry techniques.

4.6 **Systematic review- microbiome in Oesophago-gastric cancer**

4.6.1 Introduction

Since the discovery that *Helicobacter pylori* is associated with peptic ulcer disease, there has been increasing interest in the way in which bacterial species can influence the development of intestinal disorders including cancer. In particular, research has demonstrated that bacteria within gastrointestinal tract alter the physiological and immunological functions of the gut[47].

There is however a wide variation of the gut microbiota at the genus level and higher. The number of bacteria markedly changes starting from the oral cavity, through the oesophagus and stomach having 101 per gram of contents ending up with 1012 per gram of contents in the colon and distal rectum[47]. It is therefore important to value further methodological approaches to identify bacterial species and to understand their interactions towards immune response in host.

This systematic review intends to identify the current published literature and evaluate variation in microbial diversity in patients with oesophago-gastric adenocarcinoma in comparison with healthy controls.

4.6.2 Methods

A literature search (title and abstract) was performed using Ovid Medline® (1948-2016), Embase® (1974-2018), Web of Science® and PubMed® electronic databases up to and including 11th July 2018. This search was used to identify relevant studies that had assessed the microbial diversity of patients with oesophago-gastric cancer. The search also identified studies, which had compared results to a control group of healthy individuals (either healthy volunteers or those with benign disease of the upper gastrointestinal tract).

The search was undertaken using the following terms: *microbiome* OR *microbiota* AND (*Cancer* OR *malignancy* OR *neoplasm)* AND (*Gastric* OR *Stomach* OR *Ooesophageal* OR *Oesophageal*) as well as the medical subject headings (MeSH), *microbiome* OR *microbiota AND* (*stomach neoplasms* OR *oesophageal neoplasms*). A second search was performed using the terms bacteria AND (*Cancer* OR *malignancy* OR *neoplasm)* AND (*Gastric* OR *Stomach* OR *Ooesophageal* OR *Oesophageal*) as well as MeSH terms (*stomach neoplasms* OR *oesophageal neoplasms*).

The electronic search was supplemented by a hand-search of published abstracts from relevant conference proceedings (2010 – 2018). Two reviewers (M Adam and P Patel) independently screened titles and abstracts of studies identified through the electronic search. Full texts of potentially relevant articles were retrieved. Further potentially relevant articles were identified through the searching of reference lists of relevant studies. Studies were included if they reported differences in microbial composition of oesophageal or gastric mucosa from the patients with oesophagitis, chronic gastritis, intestinal metaplasia and/or healthy tissues compared to oesophago-gastric cancer. Exclusion criteria were:

studies primarily analysing *H. pylori* only rather than other microbiome present; in vitro cell line studies; animal studies; studies without a control group; conference abstracts and studies not published in the English language. Studies that reported the same patient population were also excluded, except for the most recent or complete publication.

The two reviewers independently extracted data from selected studies including primary author; year of publication; number and types of significantly diverse microbial communities in cancer and control groups especially in *Helicobacter* dominant groups. The primary outcome measure was the identification of differences in the composition and diversity of bacteria found in cancer and control subjects. The hypothesis under investigation was whether the diversity of bacteria would be significantly different between cancer and control samples.

4.6.3 Results

Six studies met the inclusion criteria (Figure 1). Details of included studies are provided in Table 2. Given the limited available data, no meta-analysis was undertaken.

Author and year of Publication	Analytical Method	Cancer site	N Cancer	Control	N Control	Bacteria Upregulated
Johan Dicksved **2009**	T-RFLP	Gastric	6	NGT NUD	5	*Streptococcus,Lactobacillus, Veillonella, Prevotella, Nesisseria flava, Hemophilus*
Francisco **2014**	G3 PhyloChip	Gastric	5	IM NAG	5 5	Lactobacillus *coleohominis,* Lachnospiraceae
Chang Soo Eun **2014**	454 GS FLX Titanium	Gastric	11	IM CSG	10 10	Streptococcus, Lactobacillus,
Yalda **2014**	MALDI-TOF	Gastric	8	NUD PUD	185 22	Acinetobacter baumannii, Klebsiella pneumonia
Hyun **2016**	Pyrosequencing	Gastric	15	NOTH	16	Staphylococcus, Neisseria flava, Klebsiella pneumonia
Daffolyn **2017**	Illumina MiSeq	Oesophagus	19	NSC BO	20 23	*Lactobacillus*

Table 2:Summary of a systematic review of the literature describing the bacteria upregulated in cancer versus control patients, NGT: normal gastric tissue; NUD: non ulcer dyspepsia; IM: intestinal metaplasia; CSG: chronic superficial gastritis; PUD: peptic ulcer disease; NOTH: normal oesophageal tissue; NAG: non-atrophic gastritis; NSC: normal squamous control; BO: Barrett's Oesophagus.

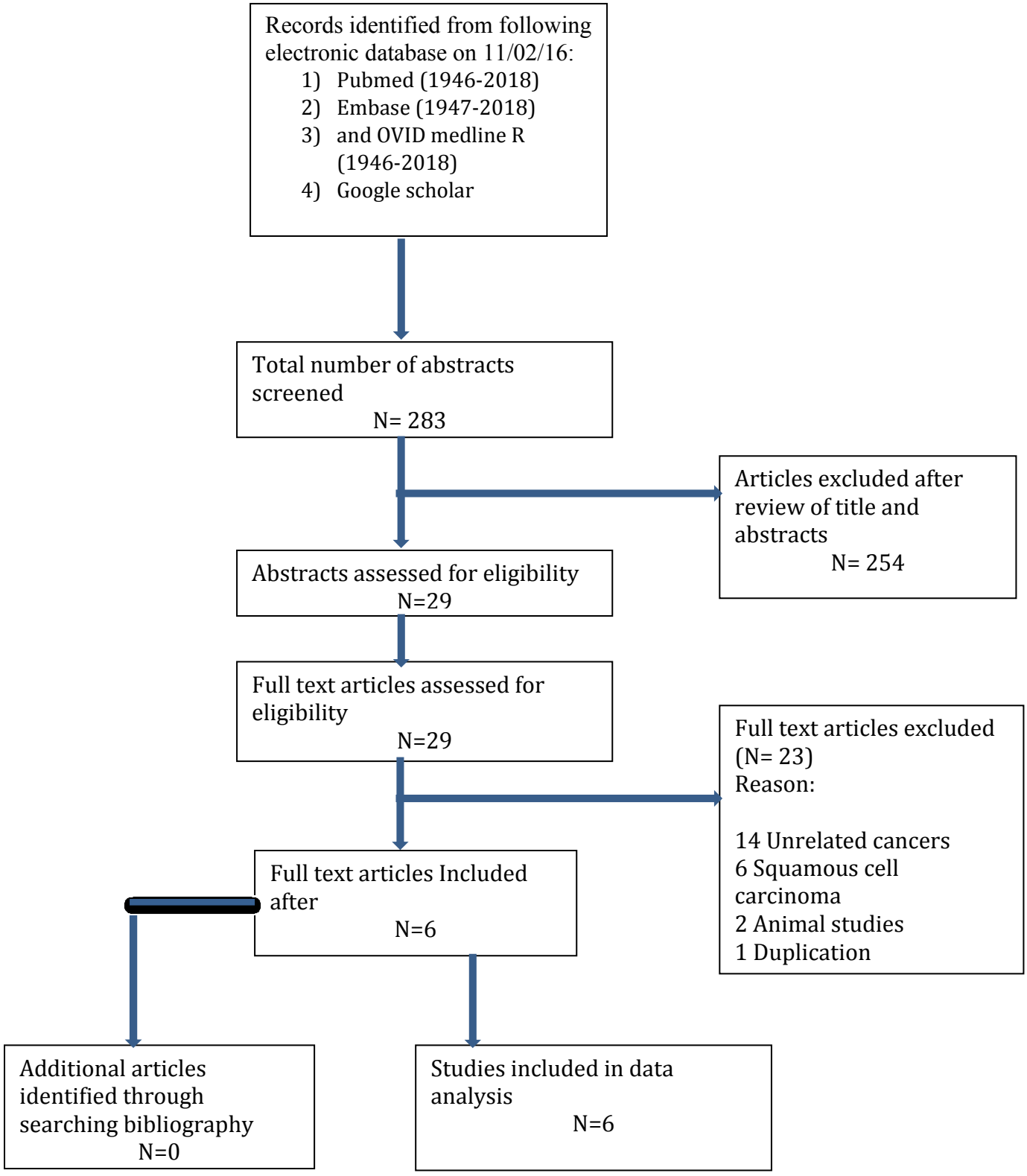

Figure 1:Systematic search and selection strategy.

In 2009, one of the earliest studies conducted by Dicksved *et al.* assessing the composition of gastric microbiota of patients with gastric cancer versus control patients with normal healthy mucosa, detecting several genera using terminal resection fragment length polymorphism (T-RFLP) in combination with sequencing of 16S rRNA genes[48]. The results demonstrated an upregulation of Firmicutes in the cancer cohort, although there were only six patients in this study. Within the subclassification of the phylum Firmicutes the vast majority of genera were Streptococcus, Lactobacillus and Veillonella. Within Phylum Bacteroidetes, Prevotella was mostly found. However, all five classes of Proteobacteria phylum were found, with Neisseria and Haemophilus being mainly upregulated.

Using G3 PhyloChip as a more sensitive approach, Francisco *et al.* has identified a higher diversity of bacterial genus level but lower number of phyla when comparing patients with gastric cancer versus non-atrophic gastritis and atrophic metaplasia[49]. At a genus level Lactobacillus coleohominis and lachnospiraceae were found in higher abundance in patients with gastric cancer compared to patients with benign disease of the upper gastrointestinal tract. Haemophilus parainfluenzae and Veillonella ratti were found to be higher in cancer patients, which has only been reported previously in the urogenital tract. This suggests that using an improved sensitive technique; various undescribed bacteria may also be able to colonize the gastric mucosa.

Soo Eun *et al.* investigated the microbiota of ten patients with chronic gastritis, ten with intestinal metaplasia and eleven with gastric adenocarcinoma using 454 GS FLX Titanium sequencer[50]. In the gastric cancer group, there was a significant increase in the Bacilli genera containing Streptococci and Lactobacilli, in comparison to *H pylori*. However, *H pylori* was

predominant in the intestinal metaplasia and gastritis groups compared to gastric cancer indicating its effects may not be overestimated in established invasive adenocarcinoma.

Yalda et al. explored the impact of *helicobacter pylori* on the diversity of human microbiota on gastric disease patients including patients with nonulcer dyspepsia (NUD), peptic ulcer disease (PUD) and gastric cancer[51]. Instead of only relying on 16S rRNA sequencing, they developed a method to employ MALDI-TOF mass spectrometry on selected colonies that have been grown from stomach biopsies on plates and picked according to their morphological features. Comparing *H pylori* positive and negative patients, data demonstrated no significant variations between groups. There was an increase of Klebsiella pneumonia and Acinetobacter baumannii in gastric cancer patients but the results were considered with caution due to the relatively low numbers of gastric cancer (n=8) compared to NUD (n=185) and PUD (n=22) groups.

Hyun *et al.* showed no significant findings between *H pylori* positive and negative patients, however an increase of staphylococcus, Klebsiella pneumonia and Neisseria flava in cancer groups versus controls[52].

In terms of assessing the microbial diversity in oesophageal adenocarcinoma, Daffolyn *et al.* has demonstrated a decrease in diversity from cancer tissues versus normal squamous controls and Barrett's oesophagus[53]. Samples collected using the Cytosponge™ device have shown a higher level of microbial taxa when compared with tissue biopsies. Lactobacillus fermentum was enriched in the microenvironment of cancer patients, showing that the alteration in microbial communities may enable the Cytosponge™ device may be applicable for early disease detection.

4.6.4 Discussion

This is the first systematic literature search and review of the characteristics of the upper gastrointestinal microbiome in oesophago-gastric cancer patients. Analysis of microbial diversity was shown to be feasible and reliable through different sequencing methods. Limited studies have shown the significance of gene sequencing to understand the functionality of the microbiome in its entirety, as well as its individual colonies.

Whilst there is a consistent increase in the identification of new microbes, particularly in research focusing on the gut microbiome, recent advances in technology have uncovered a large number of uncultured species. Despite recent advances in the technology available to analyse the microbiome, one of the biggest challenges faced by researchers is the 'dark matter' of microbiome research, often termed the 'unculturables'. This refers to the large number of microbial species that are too fastidious to be grown inside the laboratory, outside of their normal habitats, in addition to the unknown species we're yet to identify. As this is an emerging field, bacterial identification continues to be challenging as established repositories are continuously updated. Further effort is needed to identify all these unculturable and unknown microbes using culture-independent techniques that target the genetic information of these communities instead of trying to grow them in the laboratory.

The variations in microbial diversity between oesophago-gastric cancer and healthy patients provided further evidence that the stomach microbiome has the potential to produce VOCs. The use of breath analysis as non-invasive tests for the diagnosis of oesophago-gastric cancer will not be widely adopted by the clinical community unless it is possible to explain the underlying mechanisms responsible for the observed changes. Establishing the effects of microbial diversity will help to introduce these non-invasive methods for stratification of

oesophago-gastric cancer risk into clinical practice, and thereby allow for streamlining the use of endoscopy services to those at greatest risk of this disease. It is hoped this would lead to an increase rate of diagnosis at early-stages of disease before symptoms of more advanced disease such as dysphagia develop. This would lead to more patients being treated with curative intent and thereby improve survival.

The ultimate aim is to provide the scientific mechanistic pathway for a point-of-care VOC based test to differentiate oesophago-gastric cancer from healthy patients. The potential benefits for patients are: earlier diagnosis of oesophago-gastric cancer; better outcomes as a consequence of appropriate treatment, and; better patient satisfaction because of the non-invasive nature of the test.

4.7 **Thesis aims**

Research questions:

i. Does the upper gastrointestinal microbiome contribute to VOC production in esophagogastric cancer?

ii. Can the tumour associated microbiome be exploited for the purpose of augmented VOC production?

Primary aim: To investigate the role of the tumour associated microbiome in VOC production in esophagogastric cancer.

Secondary aims:

i. To characterise the upper gastrointestinal microbiome associated with oesophago-gastric adenocarcinoma

ii. To define VOC release from oesophago-gastric tumour and their associated microbiome in both *in vitro* and *in vivo* experiments

iii. To optimise a methodology for augmented microbiome mediated VOC production *in vitro*

iv. To conduct the first human trials of the proposed microbiome mediated augmented breath test in oesophago-gastric cancer patients.

5 MASS-SPECTROMETRY ANALYSIS OF MIXED-BREATH, ISOLATED-BRONCHIAL-BREATH, AND GASTRIC-ENDOLUMINAL-AIR VOLATILE FATTY ACIDS IN OESOPHAGO-GASTRIC CANCER

5.1 Background

The chemical analysis of volatile organic compounds (VOCs) in humans is a rapidly evolving field that has the potential to contribute to the non-invasive detection of multiple disease states. A recent systematic review on the diagnostic accuracy of VOC-based exhaled breath tests showed their potential for non-invasive cancer detection[54]. Previous studies have reported higher concentrations of specific VOCs, including volatile fatty acids (VFAs) and phenol, within the exhaled breath, gastric content and urine of patients with ooesophago-gastric cancer[19-24]. Whilst several studies have suggested a role for these volatile compounds in important regulatory processes in oesophago-gastric cancer[55-57], many of the biochemical pathways relating to their origin in humans are as yet unknown.

Notwithstanding, it is postulated that deregulated production of specific VOCs occurs directly from cancer tissues. These may pass in to the systemic circulation with subsequent partition across the alveolar-capillary barrier. Alternatively, VOCs may be released directly by the mucosa of the aerodigestive tract[58,59]. Hence, targeted quantification of these compounds within the headspace of oesophago-gastric tissue and isolated body compartments may prove helpful in determining the origin and mechanisms of release of these compounds. However, there remain potential VOC contributions from other endogenous sources including healthy tissues within the gastrointestinal tract and systemic effects from other organ systems.

Given the inherent complexity of biological matrices, the identification, characterisation and accurate quantification of VOCs remains challenging. In recent years, there have been

technological advances in gas phase analytical techniques allowing one to measure VOCs emitted from the headspace of biofluids and histological specimens with accuracy at levels to parts-per-trillion by volume (*pptv*). In particular, mass spectrometry techniques including Proton-Transfer-Reaction Time-of-Flight Mass Spectrometry (PTR-TOF-MS) and Gas Chromatography Mass Spectrometry (GC-MS) have been widely utilised for VOC detection in human studies[60-63]. PTR-TOF-MS is notable its ability to perform real-time VOC quantification measurements and analysis of a full mass spectrum within a fraction of a second as well as separation and identification of isobaric compounds.

The purpose of this study is to investigate production of targeted VFAs and phenol in oesophago-gastric cancer through analysis of *ex vivo* headspace above tissues and *in vivo* within different anatomical compartments including mixed breath, isolated bronchial breath and gastric endoluminal air. Determining the relative abundance of VOCs within these compartments may provide a clearer understanding of their source of origin and association with oesophago-gastric cancer.

5.2 **Ex-vivo tissue headspace analysis**

5.2.1 Aims

This experiment aimed to establish the presence of VOCs within non-derivatised tissue samples from oesophago-gastric cancer patients and control subjects.

5.2.2 Methods

5.2.2.1 *Patients selection and sample collection*

Patients for this study were recruited through the Gastrointestinal Endoscopy Unit at St Mary's Hospital (London). Patients are referred for an Oesophago-Gastro-Duodenoscopy (OGD) investigation after clinical evaluation from their doctor. OGD referral forms were screened for suitable patients for recruitment into this study. Patients were requested to provide a comprehensive medical history, smoking status and alcohol intake history. The exclusion criteria included patients with known liver disease, small bowel/colonic conditions or any other form of cancer. Patients with a biopsy-confirmed invasive oesophago-gastric adenocarcinoma were included in the cancer cohort. Patients with biopsy confirmed non-cancer conditions of the upper gastrointestinal tract (e.g., esophagitis, gastritis and peptic ulcer disease) were included in the positive control cohort. Patients who had a normal OGD test and a negative rapid urease test (for the *H.pylori* bacterium) were included in the 'healthy upper gastrointestinal tract' cohort. Local ethics committee approval was granted

for this study (Ref:12/WA/0196) and written informed consent was obtained from all patients prior to enrolment in the study.

In addition to demographic details, the following meta-data were collected for all subjects: reason for initial referral; past medical history; current medications; smoking and alcohol status; endoscopy findings; histopathology findings; cancer stage (if cancer confirmed); H-Pylori status; time since last oral intake.

After washing away any overlying organic material or food residue, tissue samples from the upper gastrointestinal tract were retrieved at the time of endoscopy using cold biopsy forceps (2.8mm EndoJaw, Olympus, UK). Biopsies were taken from diseased portions of the stomach/oesophagus (cancer or benign disease) and from healthy mucosa from the body of the stomach in patients without obvious pathology on visual inspection of the upper gastrointestinal tract. Tissue samples were extracted from the cold biopsy forceps in a sterile manner and immediately snap frozen in liquid nitrogen. All samples were logged and stored at -80^{o}C until analysis.

5.2.2.2 *Sample analysis*

Analysis of VOCs was conducted using PTR-TOF-MS (PTR-TOF 1000, Ionicon Analytik GmbH, Innsbruck, Austria). The PTR-TOF-MS technique has been extensively described elsewhere [64] allowing for a brief overview to be provided herein. The principal components of the PTR-TOF-MS instrument include an ion source coupled with a drift tube and high mass resolution, orthogonal acceleration, reflection time-of-flight mass spectrometer[64-66], Precursor ions, H_3O^+ and NO^+, produced within the ion source are allowed to react with a sample gas that is introduced into the drift tube. Characteristic protonated product ions are

formed from the proton transfer reactions that occur between product ions and trace gases within the sample. Product ions are separated according to their *m/z* ratio in the time-of-flight and detected by a sensitive ion detector and associated data acquisition electronics. In this way VOCs may be accurately detected and quantified in real-time at sub parts-per-billion by-volume (*ppbv*) levels.

Drift tube conditions were: temperature 110°C, pressure 2.30 mbar and voltage 350 V, resulting in an E/N of 84 Td (1 Townsend = 10^{-17} V cm^2). Sampling was carried out by means of a heated (110°C) PEEK tubing and the inlet flow was set at 130 sccm. Headspace analysis was performed by piercing the septum of the vial with a sterile needle attached to the PTR-TOF-MS sample inlet. A second perforation in the septum was created to allow clean air to be flushed in to the vial at a rate of 40ml/min. VOC concentrations could be calculated according to:

(1) $$C_x = (S_x / Tr_x) / [(S_{21} \times k_{21} / tr_{19} + S_{39} \times k_{39} / tr_{37}) \times t_{react} \times N]$$

(2) $$C_x = (S_x / Tr_x) / [(S_{30} \times k_{30} / tr_{30}) \times t_{react} \times N]$$

Equations (1) and (2) can be used to calculate the concentration of compound X (C_x expressed as volume mixing ratio), using H_3O^+ and NO^+ as primary ions, respectively. S_x represents the ion signal of compound X. When using H_3O^+ as primary ion (equation 1), S_{19} and S_{37} represent the hydronium and protonated water dimer ion $(H_2O)_2H^+$. To overcome primary ion signal saturation, hydronium ion and cluster signals were extrapolated from the corresponding ^{18}O isotopologues, at *m/z* 21.02 and 39.03 Th, respectively. When employing NO^+ as primary ion (equation 2) S_{30} represents the nitrosonium ion. To overcome mass peak saturation, the NO^+ signal was extrapolated from the corresponding ^{15}N isotopologue at m/z

30.99 Th. Number density (N) and residence time in the drift tube (t_{react}) was calculated applying a well-established theory[67.] Reaction rate constants of compound X with the respective primary ions (k_{19}, k_{37} and k_{30}) were calculated using Su's parametrisation approach[68]. This allows for non-thermal conditions typical of commercial PTR-MS instruments, permitting more accurate quantification[69]. All raw signals were modified correcting for the detector-specific transmission factors (tr_x, tr_{19}, tr_{37}, tr_{30}). The transmission curve was obtained by measuring a certified mixture of aromatic hydrocarbons, all at the concentration of 100 ppbV (TO-14 Aromatics Subset Mix, Thames Restek Saunderton UK). The transmission values were established by interpolating the measured intensity values of reference mass peaks by means of a natural spline.

5.2.2.3 *Method optimisation*

Additional experiments were undertaken to evaluate the effect of freezing and to determine the optimal time interval sample headspace necessary prior to analysis. Twenty healthy volunteers (with no comorbidities) were recruited and requested to provide tissue samples at the time of routine OGD. Two tissue samples were obtained from each patient and placed in two different vials. One sample was snap frozen using liquid nitrogen and the other was directly analysed using the H_3O^+ precursor ion of the PTR-TOF-MS to assess the effect of freezing between samples. Septum was pierced at 5 different time points to allow tissue production of volatile compounds. The results demonstrated some differences between both batches this may be due to volatile loss during freezing of tissue samples but due to practical reasons frozen method was used for the main experiment performed. However, there were subtle differences between different times points.

The influence of length of time headspace was allowed to 'develop' and the effect of tissue freezing were initially investigated in order to establish an appropriate methodology. Analysis of duplicate gastric biopsy samples from healthy patients were analysed at set intervals using the H_3O^+ precursor ion of the PTR-TOF-MS. With the exception of acetone, target VOCs measured at all time points within the headspace above fresh (unfrozen) biopsy samples were detected at either equivalent of significantly greater levels compared to biopsy samples that had been stored after freezing (full details of results of method development experiments are provided in table 3 below). In comparison acetone levels were significantly higher at all time points when detected above sample that had previously been frozen. Irrespective of freezing, headspace VOCs tended to exhibit a decay when measured over serial time points with highest levels observed at either 0.0 or 0.5 hours.

Based on these findings and the logistical challenges of performing regular analysis of fresh tissue samples, all subsequent experiments were conducted using samples that had been frozen after retrieval and then subsequently thawed to permit headspace analysis after a time interval of 0.5 hours.

		Frozen		Unfrozen		*P**
	Time (mins)	Median (ppbv)	IQR	Median (ppbv)	IQR	
Acetaldehyde	0.0	3.73[1]	[2.32-7.74]	3.56	[2.73-6.54]	0.837
	0.5	1.98	[1.34-2.84]	2.92	[0.62-4.36]	0.806
	1.0	1.45	[1.18-1.82]	3.83	[0.67-4.21]	0.244
	2.0	1.69	[1.17-1.76]	4.33	[2.49-5.91]	0.007
	4.0	1.61	[1.43-2.11]	2.88	[1.88-6.14]	0.039
Acetone	0.0	10.75	[9.32-37.47]	5.97[1]	[4.90-8.17]	<0.001
	0.5	13.52	[11.10-20.15]	3.78	[3.18-4.93]	<0.001
	1.0	10.18	[8.52-17.76]	3.22	[2.42-4.64]	<0.001
	2.0	9.32	[6.38-13.03]	3.94	[3.05-5.82]	0.012
	4.0	8.79	[7.01-13.30]	4.39	[3.64-4.87]	0.002
Acetic acid	0.0	2.24[2]	[1.26-4.18]	7.86[1]	[6.45-10.50]	<0.001
	0.5	6.16	[4.68-8.39]	6.72	[4.64-7.98]	1.000
	1.0	5.61	[4.84-6.79]	5.10	[3.80-6.92]	0.317
	2.0	5.10	[3.46-6.28]	5.71	[4.59-7.92]	0.523
	4.0	5.36	[4.95-8.70]	3.72	[2.73-9.21]	0.077
Butyric acid	0.0	0.53[1]	[0.31-0.86]	0.99	[0.73-1.38]	<0.001
	0.5	1.14	[0.96-1.78]	0.98	[0.64-1.36]	0.175
	1.0	1.02	[0.62-1.32]	0.74	[0.42-0.89]	0.072
	2.0	0.72	[0.61-0.96]	0.76	[0.51-1.31]	0.569
	4.0	0.72	[0.66-1.11]	0.62	[0.38-0.94]	0.319
Pentanoic acid	0.0	0.05[1]	[0.02-0.07]	0.44	[0.40-0.58]	<0.001
	0.5	0.17	[0.13-0.19]	0.40	[0.25-0.50]	<0.001
	1.0	0.12	[0.10-0.15]	0.37	[0.24-0.46]	<0.001
	2.0	0.08	[0.07-0.09]	0.39	[0.31-0.81]	<0.001
	4.0	0.08	[0.07-0.11]	0.50	[0.30-0.84]	<0.001
Hexanoic acid	0.0	0.13[1]	[0.04-0.17]	0.52[3]	[0.38-0.56]	<0.001
	0.5	0.22	[0.18-0.28]	0.41	[0.29-0.48]	<0.001
	1.0	0.18	[0.14-0.20]	0.34	[0.23-0.43]	<0.001
	2.0	0.11	[0.08-0.13]	0.37	[0.30-0.65]	<0.001
	4.0	0.10	[0.09-0.13]	0.35	[0.26-0.60]	<0.001

Table 3:Results of experiments examining the effect tissue freezing and sample time point of the detection of selected volatile organic compounds in the headspace above healthy gastric tissue samples. IQR, interquartile range. **P*-value; frozen vs. unfrozen, Mann Whitney U test. Friedman test comparing volatile organic compound concentrations detected at each time point 0.0 to 4.0 hours: [1]*P*<0.001, [2]*P*=0.001, [3]*P*=0.050.

Once the optimum methodology had been established, batches of frozen tissue biopsies were left to thaw at room temperature. Biopsies were placed in a 20mL screw-capped vial (Thermo Scientific, Hemel Hempstead UK), which was allowed to equilibrate at room temperature for 30 minutes. Selected VOCs were analysed using the H_3O^+ precursor ion of

PTR-TOF-MS (PTR-TOF 1000, Ionicon Analytik GmbH, Innsbruck, Austria). Below Figure 3 Illustrates the method of analysis using PTR-ToF-MS.

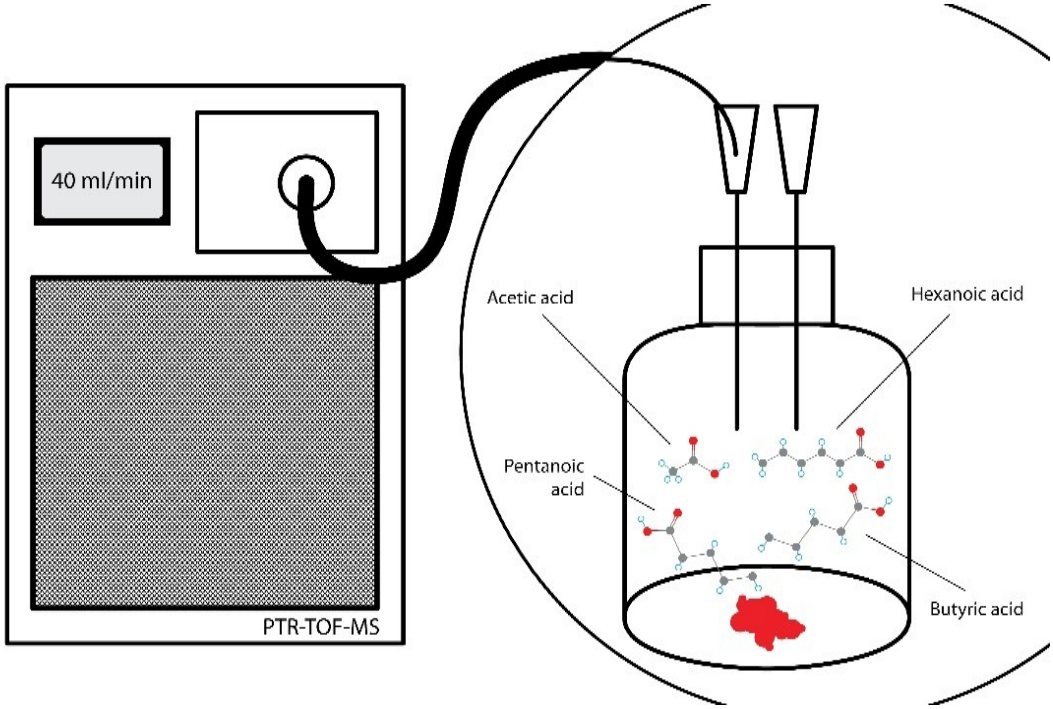

Figure 2:*Ex vivo* headspace analysis with PTR-ToF-MS.

5.2.2.4 *Quality control*

During the described experiments, a series of quality checks were conducted on the PTR-TOF-MS daily. Impurities with the two ionization modes were O_2^+ (<2%) for H_3O^+ as primary ion and NO_2^+ (<3%) for NO^+ as primary ion, respectively. Quantitation accuracy was within ±10% of a certified standard, represented by a Trace Source™ benzene permeation tube (Kin-Tek Analytical Inc., La Marque TX). The repeatability of fragmentation patterns with NO^+ and H_3O^+ as primary ions was assessed by measuring the ratio between reference mass peaks with given standard compounds. For NO^+, we used the ratio between peaks *m/z* 71 and 43: these represent the quasi-molecular and the most representative fragment for

butanal, as obtained from a permeation tube standard. For H_3O^+, the ratio between peaks *m/z* 89 and 71 were used to represent the quasi-molecular and the most representative fragment for butyric acid, as obtained from a permeation tube standard. The values measured on the different days were within ±2% of the mean. Whenever required, we optimized the voltage of the microchannel plate and the mass resolution (>1,500 m/□m), using *m/z* 89 (butyric acid with H_3O^+) as reference peak.

5.2.2.5 *In-situ analysis of tumour headspace*

In a single patient, direct headspace analysis of a gastric tumour and adjacent 'normal' mucosa was performed immediately after surgical removal of the whole stomach. Following resection, the stomach was immediately opened along the greater curve to exposure the entire gastric mucosa. A sterile polystyrene sample container (60mL) was modified to permit the passage of the PTR-TOF-MS sample line through its base and was placed over the tumour and the headspace was analysed for 60 seconds. Headspace above adjacent gastric mucosa that was macroscopically uninvolved by tumour was subsequently undertaken.

5.2.2.6 *Statistical analysis*

Statistical analysis was performed using IBM SPSS statistics 22 (SPSS Inc., Chicago, IL). The Kruskal-Wallis test was used to compare the measured concentrations of VOCs between the three groups. The Mann-Whitney U test was also applied to compare the measured concentrations between the oesophago-gastric Cancer and healthy upper gastrointestinal tract control groups. A *p*-value ≤ 0.05 was taken as the level to indicate statistical significance. A total of 5 Volatile Organic Compounds released from the headspace of tissue were analysed using PTR-ToF-MS. These compounds were selected based on previous VOC

studies in oesophago-gastric cancer[19-22]. The analytical information, including chemical formula, precursor ions, *m/z* ratio and characteristic product ions are included in Table 4.

Compounds	Molecular formula	*m/z*	Characteristic product ions
Acetone	C_3H_6O	59.049	$C_3H_6OH^+$
Acetic acid	$C_2H_4O_2$	61.028	$C_2H_4O_2\ H^+$
Butyric acid	$C_4H_8O_2$	89.060	$C_4H_8O_2H^+$
Pentanoic acid	$C_5H_{10}O_2$	103.075	$C_5H_{10}O_2H^+$
Hexanoic acid	$C_6H_{12}O_2$	117.091	$C_6H_{12}O_2H^+$
Phenol	C_6H_6O	95.049	$C_6H_6O^+$
Acetaldehyde	C_2H_4O	45.033	CH_3CHOH^+

Table 4:Summary of analytical information for compounds detected and quantified by PTR-ToF-MS using the H_3O^+ precursor ion.

5.2.3 Results

Biopsy samples were collected from 45 patients with oesophago-gastric cancer (31 male, 72 ± 14yrs), 19 positive control (10 male, 54±12yrs) and 64 healthy controls (36 male, 57±17yrs). The average weight of biopsy tissue samples was 5.5±4.3mg.

Target VOCs were significantly increased above oesophago-gastric cancer tissue biopsies compared to healthy controls (Table 5). Importantly, the same VOCs were also increased, albeit to a lesser extent, in the headspace of upper gastrointestinal mucosa affected by benign inflammatory conditions. Average biopsy weight was 5.5±4.3mg and did not correlate with any measured VOC concentration (R^2<0.010, P>0.086).

Fatty acids were all significantly increased in oesophago-gastric cancer tissue in comparison to healthy controls. The measured acetic acid concentrations of the cancer tissue were 10.7*ppbv* in comparison to the positive control and healthy groups, which were 6.9*ppbv* and 4.9*ppbv,* respectively. (Figure 3).

	Cancer Cohort N=45	**Positive Control N=19**	**Healthy controls N=64**	***P*[1]**
Acetone	42.4 [10.8-365.1]	22.7 [10.1-42.9]	12.7 [6.0-26.2]	<0.001
Acetic acid	10.7 [4.7-26.0]	6.9 [5.3-16.6]	4.6 [0.1-8.5]	0.001
Butyric acid	2.2 [0.7-5.9]	1.1 [0.3-1.8]	0.7 [0.4-1.6]	0.004
Pentanoic acid	0.5 [0.2-0.7]	0.4 [0.3-0.6]	0.3 [0.2-0.5]	0.029
Hexanoic acid	0.6 [0.2-2.8]	0.6 [0.3-1.0]	0.2 [0.1-0.7]	0.033

Values are presented as median and interquartile range. [1]Kruskal–Wallis test

Table 5:Headspace concentrations (*ppbv*) of VOCs significantly increased in esophago-gastric cancer tissue, Values are presented as median and interquartile range. [1]Kruskal–Wallis test.

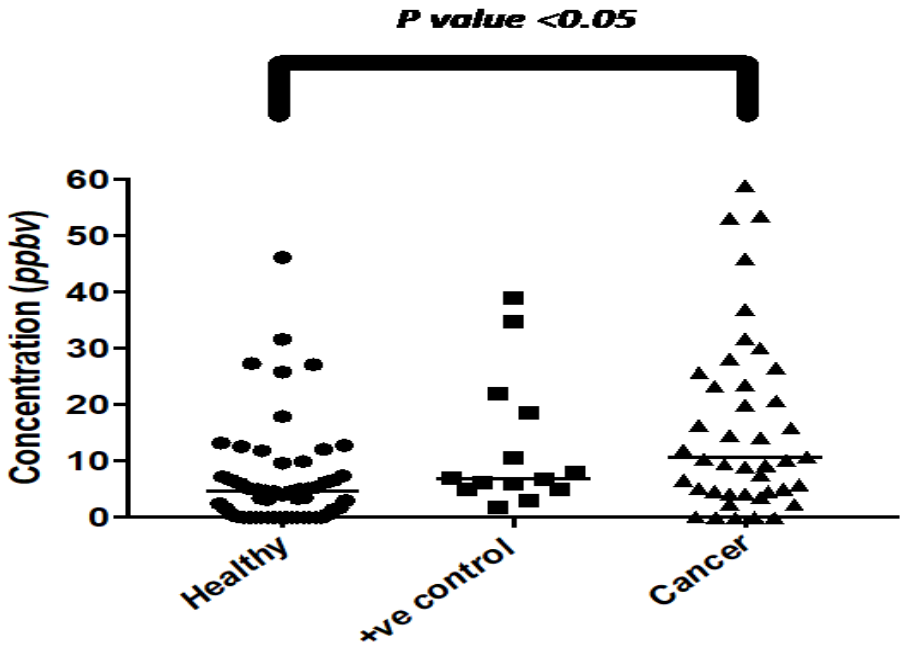

Figure 3:Scatter plots of the median concentrations (in *ppbv*) of acetic acid from patients with oesophago-gastric cancer, positive (+ve) controls and the healthy upper GI tract cohort.

The measured butyric acid concentrations of the cancer tissue were 2.2*ppbv* in comparison to positive control and healthy groups, which were 1.1*ppbv* and 0.7*ppbv,* respectively (Figure 4).

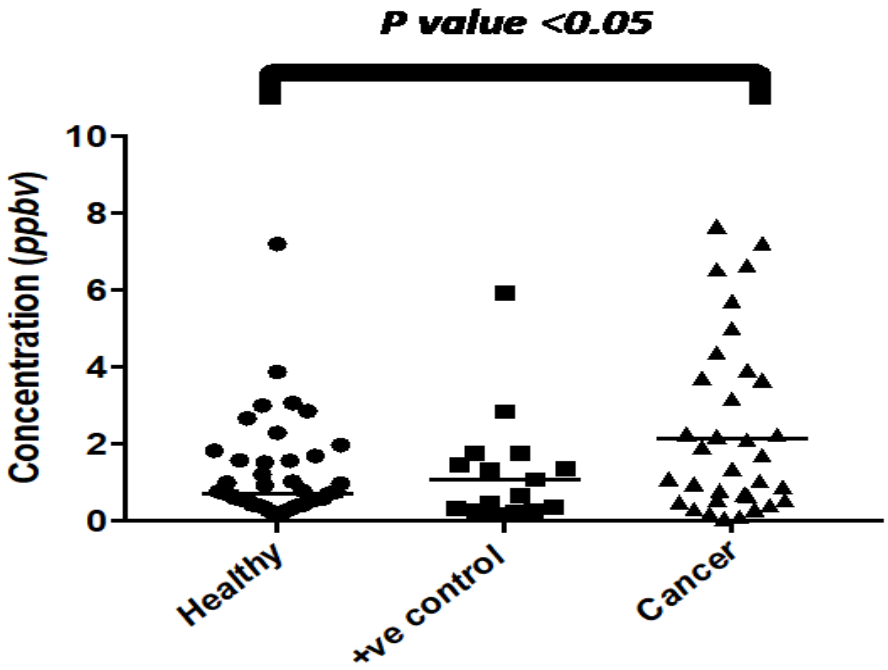

Figure 4:Scatter plots of the median concentrations (in *ppbv*) of butyric acid from patients with oesophago-gastric cancer, positive (+ve) controls and the healthy upper GI tract cohort.

The measured Pentanoic acid concentrations of the cancer tissue were 0.5*ppbv* in comparison to positive control and healthy groups, which were 0.4*ppbv* and 0.3*ppbv,* respectively. The measured hexanoic acid concentrations of the cancer tissue were 0.6*ppbv* in comparison to positive control and healthy groups, which were 0.6*ppbv* and 0.2*ppbv,* respectively (Figure 5).

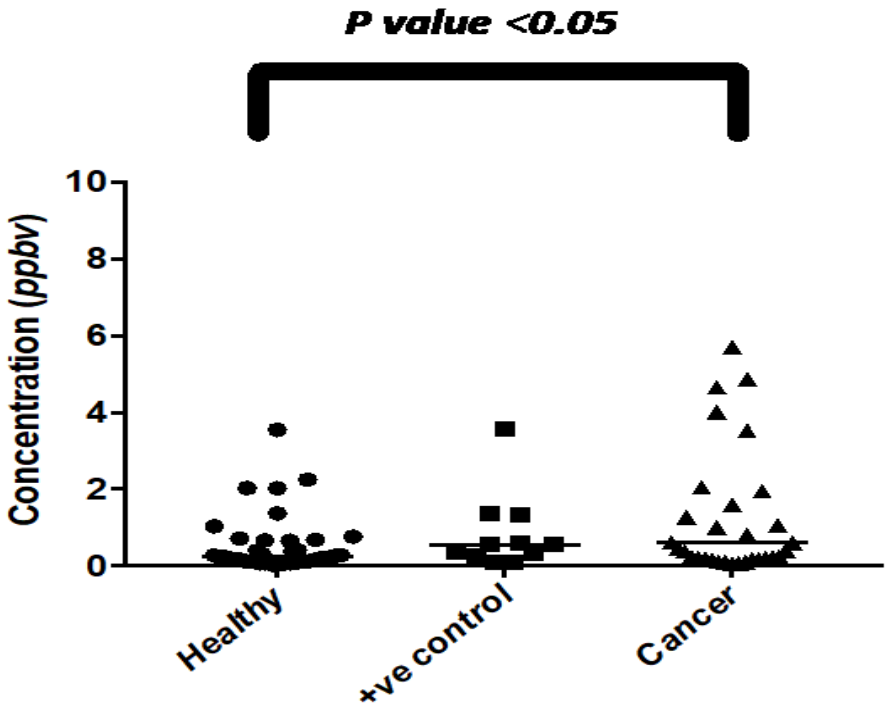

Figure 5:Scatter plots of the median concentrations (in *ppbv*) of hexanoic acid from patients with oesophago-gastric cancer, positive (+ve) controls and the healthy upper GI tract cohort.

The measured Phenol concentrations of the cancer tissue were 2.1*ppbv* in comparison to positive control and healthy groups, which were 1.2*ppbv* and 0.6*ppbv,* respectively. (Figure 6).

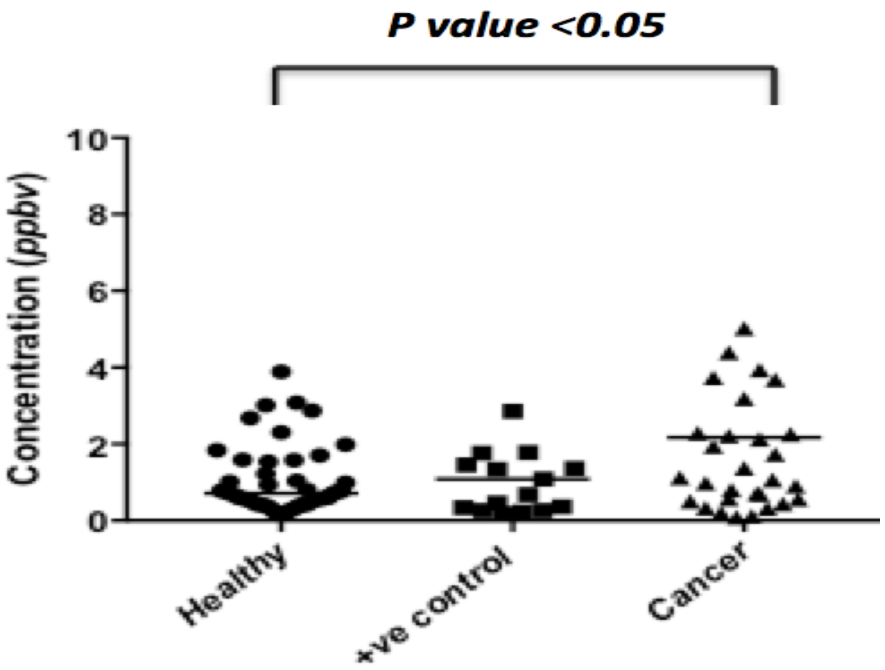

Figure 6:Scatter plots of the median concentrations (in *ppbv*) of Phenol from patients with oesophago-gastric cancer, positive (+ve) controls and the healthy upper GI tract cohort.

Other compounds that were analysed including (ethanol, acetaldehyde, Isoprene and ammonia) were not significantly altered in cancer. Tissues affected by benign inflammatory pathologies (i.e., positive control group) expressed an intermediate phenotype with acetone and fatty acid VOC concentrations that were higher than in healthy tissue but lower than those in cancer.

5.2.3.1 *In-situ analysis of tumour headspace*

Direct sampling PTR-ToF-MS analysis of the headspace of a gastric cancer immediately following surgical resection of the whole stomach was performed in a single patient (Table 6). Acetone (795.3 vs. 388.8*ppbv*), acetic acid (29.0 vs. 18.1*ppbv*), butyric acid (2.8 vs 1.8*ppbv*), pentanoic acid (1.1 vs 0.8*ppbv*) and hexanoic acid (1.7 vs 1.0*ppbv*) were observed at higher concentrations within the *in-situ* headspace above the tumour compared to macroscopically normal adjacent gastric mucosa, as shown in Figure 7.

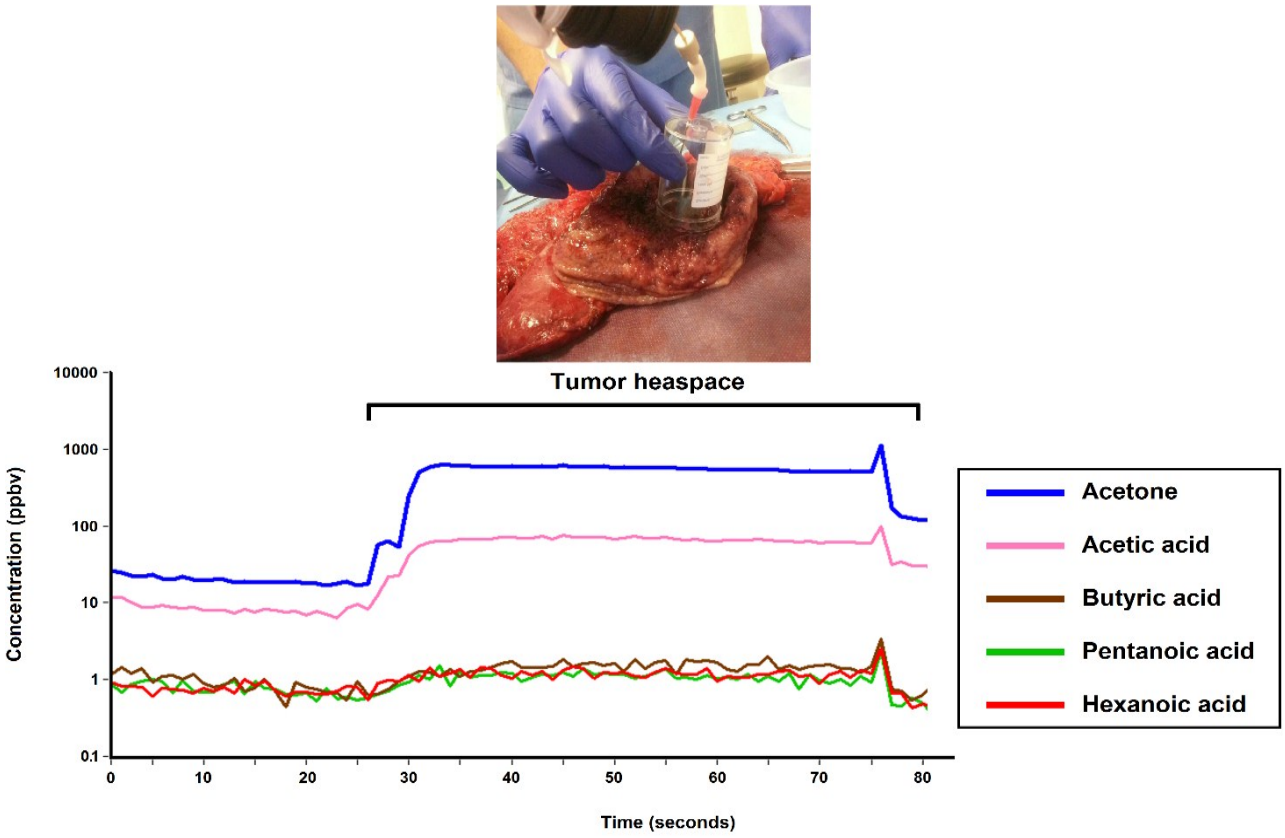

Figure 7:Direct PTR-ToF-MS mass spectrum analysis of acetone and short chain volatile fatty acids within the headspace of cancer & healthy tissue regions of a surgically removed stomach.

The VOC concentrations measured in a single patient through direct sampling analysis were appreciably higher than in tissue biopsies experiments. This may reflect the significantly greater tissue biomass above which measurements were taken or potentially analysis of the tissue being undertaken while fresh (i.e., immediately after resection of the whole stomach).

	Adjacent 'normal' mucosa (ppbv)	Cancer associated mucosa (ppbv)
Acetaldehyde	23.6	99.3
Acetone	388.8	795.3
Acetic acid	18.2	29.0
Butyric acid	1.8	2.4
Pentanoic acid	0.8	1.1
Henxanoic acid	1.0	1.7
Butanal	1.2	1.9
Pentanal	1.0	1.4
Phenol	0.6	1.1
Hexanal	1.3	2.3
Methyl phenol	0.9	0.8
Heptanal	0.9	1.4
Ethyl phenol	0.4	0.5
Octanal	0.9	1.5
Nonanal	1.2	2.0
Decanal	0.9	1.4

Table 6:Result of 'in-situ' tumour analysis headspace in a single patient.

5.2.4 Discussion

Non-invasive sampling using PTR-TOF- MS have been used previously with the objective of being able to provide a sensitive analytical technique to identify disease associated VOCs. To our knowledge, this is the first study to perform a targeted online analysis of selected VOCs within oesophago-gastric cancer tissue headspace.

The experiment has established the association between cancer and dysregulation of ketosis, fatty acid and phenol metabolism[70, 71]. The presence of increased concentrations of these compounds within the headspace of oesophago-gastric cancer biopsies suggests that the tumour itself is a contributory source for these VOCs. If the tumour is responsible for unregulated production of specific VOCs, these may subsequently pass to the circulatory system from where they may be excreted through the lungs, sweat, gastrointestinal or urinary systems.

The majority of VOCs are known to be of systemic origin and the concentrations observed through excretion reflect the levels observed in blood[72]. Acetone, acetic acid, butyric acid, pentanoic acid, hexanoic acid and phenol have all previously been observed at increased concentrations across several biological surrogates in patients with oesophago-gastric cancer[19, 21-24]. Importantly, the same VOCs were also increased (albeit to a lesser extent) in the headspace of upper gastrointestinal mucosa affected by benign inflammatory conditions; this implies their production is likely to reflect alterations in metabolic pathways common to both cancer and other inflammatory pathologies.

Acetone is produced through lipolysis or from acetyl-CoA as a breakdown product of fatty acid oxidation. In our previous studies on gastric content and urine, we observed higher

concentrations in both biofluids in the cancer cohort compared to non-cancer controls[23, 24]. In this study, the measured acetone concentrations of the cancer tissue (42.4*pbv*) were significantly increased in comparison to the non-cancer controls (22.7*ppbv* and 12.7*ppbv* for the positive control and healthy groups, respectively. The observed increase in acetone could be a non-specific marker of underlying cancer status or may reflect a change in their physiological state in response to disease. Hasim *et al.* have previously reported significantly increased blood plasma acetone concentrations in patients with poorly differentiated ooesophageal cancer[73]. There is also emerging evidence that acetone and other ketone bodies may be involved in sustaining abnormal tumour growth in cancer patients[70]. Bonucelli *et al.* reported that ketones could function as chemo-attractants and stimulate the migration of epithelial cancer cells; it was demonstrated that this phenomenon directly fuels primary tumour growth and stimulates the formation of metastases without measurable increases in tumour angiogenesis[74]. Schug *et al.* also reported that acetone is utilized as a nutritional source in experiments on cancer cell lines; the authors postulated that acetone allows tumours to maintain cell growth under conditions of metabolic stress with resultant fatty acid synthesis[75].

Cancer cell proliferation requires fatty acids for the synthesis of membranes and signalling molecules[71], in this study, volatile fatty acids have again been observed to be universally upregulated in human oesophago-gastric biosamples.

Acetic acid is a metabolic intermediate for the generation of acetyl-CoA. In our previous studies on gastric content and urine, we observed higher concentrations of acetic acid in both biofluids in the cancer cohort compared to healthy controls[23, 24]. Hasim *et al.* also reported increased levels of acetate in the NMR profile of urine in patients with

oesophageal cancer compared to healthy controls[73]. From a diagnostic perspective, acetic acid chromoendoscopy is a clinical tool employed for the detection of early tumours in Barrett's oesophagus (a premalignant condition of the 0esophagus) [76]. In this technique, acetic acid causes breakage of the disulfide bonds of the glycoproteins and the unbuffered acid on the mucosal-cell surface then causes a reversible deacetylation of cellular proteins allowing for improved diagnostics of subtle cancerous lesions[77].

In this study butyric acid was found to be upregulated above tumours. Butyric acid was not a VOC selected for investigation during our studies on gastric content, urine and initial studies on exhaled breath in oesophago-gastric adenocarcinoma[22-24]. However, in our most recent multicenter study on exhaled breath analysis for oesophago-gastric cancer, butyric acid was identified as a key discriminatory VOC between cancer and controls[19]. Shi *et al.* reported that 4-Phenybutric acid promotes gastric cancer cell migration via histone deacetylase mediated HER3/HER4 upregulation[78]. Butyric acid can also be produced from periodontopathic bacteria as an extracellular metabolite and it has been implicated in the development of oral cancer[79].

Although the observed differences in pentanoic acid concentrations between the cancer and non-cancer controls were at sub-1*ppbv*, the limit of the detection for the PTR-ToF-MS is within the *pptv* range and statistical analysis revealed these results to be significant. Moreover, both pentanoic acid and hexanoic acid were principal VOCs in the exhaled breath diagnostic prediction models for oesophago-gastric cancer in the original study and subsequent multicentre validation study[21, 19]. Using TD-GC vs GC-ToF-MS, Stadler *et al.* identified hexanoic acid as a potential marker of tissue decomposition from cadavers; in gastrointestinal cancer states, tumour necrosis could produce similar effects in the

microenvironment as seen with ex-vivo decomposition[80]. Hexanoic acid has also been reported to be significantly increased in plasma samples from patients with high-grade dysplastic colonic adenomas compared to controls [81].

Phenol was utilized alongside ten other volatile organic compounds to produce a diagnostic model with an area under the ROC curve of over 0.92[21]. The conversion of tyrosine to phenol is modulated by bacteria[82]. Therefore, it is most likely that this compound would be identified within oesophago-gastric cancer tissues which maybe in direct contact with bacteria within gastric juice which is known to contain a diverse range of bacterial species[83].

Due to the size constraints of endoscopic biopsy techniques and to ensure a universal tissue retrieval technique, it was not feasible to obtain larger tissue samples. Although VOC concentrations did not correlate with biopsy weight, the in-situ experiment suggests that larger tissue volume may generate comparably higher concentrations of target VOCs. This study only sought to analyse a selected panel of target VOCs and as such may have overlooked other important compounds associated with oesophago-gastric cancer. Finally, like previous observational studies that investigated VOCs in biological samples from patients with oesophago-gastric cancer, the current study did not involve formal histological assessment of biopsies to define the precise origin of detected VOCs, which largely remain unknown.

5.2.5 Conclusion

In conclusion, the results obtained indicate that cancer-specific VOCs are directly emitted by from tissue without the need for derivatisation. The results demonstrate that oesophago-gastric cancer-specific VOCs may arise directly from tumour tissue and/or tumour associated microbiome to provide further evidence for the development VOC profiling for cancer diagnosis. The opportunity to measure VOCs within exhaled breath and other biological samples remains an attractive prospect for non-invasive disease diagnosis and monitoring. The validation of a VOC profile specific to oesophago-gastric cancer could result in much-needed clinical diagnostic tests and potentially improve patient outcomes.

5.3 **In-vivo headspace analysis**

The purpose of this study was to investigate production of targeted VOCs in oesophago-gastric cancer through analysis of *in vivo* within different anatomical compartments including mixed breath, isolated bronchial breath and gastric endoluminal air.

5.3.1 Aims

i. To develop a method to conduct endoluminal analysis to compare luminal VOCs in oesophago-gastric cancer and control patients.

5.3.2 Methods

5.3.2.1 *Patient selection*

Subjects were recruited from St Mary's Hospital, London between June 2016 and July 2017. Patients with biopsy proven oesophago-gastric junctional adenocarcinoma admitted for elective staging laparoscopy and oesophago-gastro-duodenoscopy (OGD) under general anaesthesia were enrolled to this study. Control subjects, without endoscopic or histological evidence of malignancy, were recruited from the Endoscopy Unit of St Mary's Hospital during the same time period. These patients underwent OGD under sedation or local anaesthesia. All patients were required to be fasted and to refrain from smoking for a minimum of six hours prior to breath testing. Patients with a history of previous treatment

for cancer were excluded from the study. Ethical approval was obtained through NHS Health Research Authority (Ref: 5/LO/1140).

5.3.2.2 *Breath Sampling methodology*

Prior to OGD/surgery a 500 mL 'whole breath' sample was collected using the ReCIVA breath sampler (Owlstone, Cambridge, UK)(Figure2a). The methodology for breath sampling using this ReCIVA device has been previously published [62].

In cancer patients a sample of isolated tracheal air was obtained shortly after induction of general anaesthesia and endotracheal intubation. Tracheal air (500mL) was sampled directly onto thermal desorption (TD) tubes (Markes International, Llantrisant, UK) pre-packed with 200 mg of Tenax and 100 mg of Carbograph 5, at a rate of 250mL/min. Breath was sampled from the capnography port of the ventilator circuit throughout the respiratory cycle (Figure 1). The following standardised ventilatory settings were applied 5mins prior to and for the duration of sampling: fraction of inspired oxygen 100%; respiratory rate 10 breaths per minute, and; positive end expiratory pressure. All traces of volatile anaesthetic gases were removed from the anaesthetic circuit prior to tracheal sampling to avoid their potential influence breath gas analysis. Total intravenous anaesthesia was induced and maintained using with alfentanil and propofol. (Figure 8b).

5.3.2.3 *Intraluminal sampling of in-vivo gastric headspace*

A new method was developed to sample the gastric luminal headspace through the operating channel of a flexible endoscope. After inflation of the stomach with medical air

2mm wide V-green extension infusion line (Vygon, 5 rue Adelina 95440 France Ref: 71100.20) was advanced in to the gastric lumen during OGD. The end of the sampling line was connected to a TD tubes. A sample of 500mL luminal air was obtained at a rate of 250 mL/min achieved by a precision handheld 210-1002MTX pump (SKC Ltd, Dorset, UK). The diagram below illustrates the different sampling procedures.

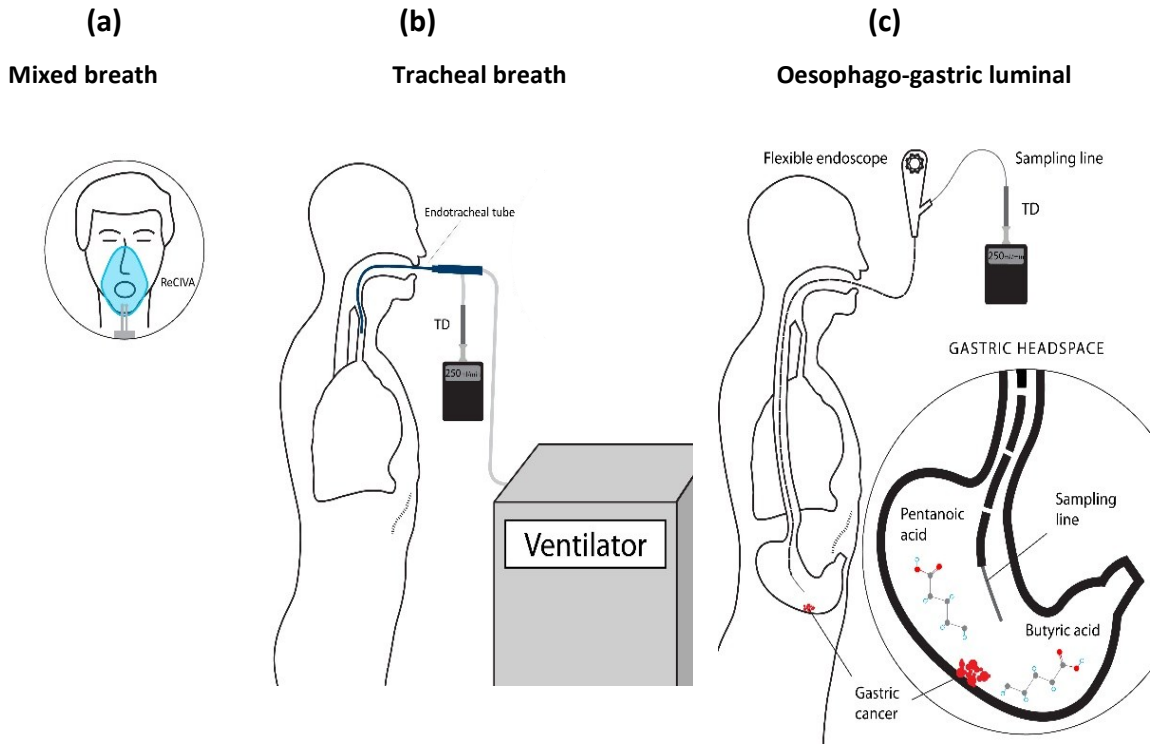

Figure 8:(a) Pre-procedure mixed breath samples collection with the ReCIVA device and (b) intra-operative sampling of the isolated bronchial breath via the endotracheal tube. (c) Sampling of the gastric endoluminal headspace via a suction channel of a standard endoscope with a custom-made catheter directly adjacent to the tumour. (TD, thermal desorption tube).

5.3.2.4 *Sample analysis by TD-GC-MS*

Samples were analysed using an Agilent 7890B GC with 5977 A MSD (Agilent Technologies, Cheshire, UK), coupled to a Markes TD-100 device (Markes International, Llantrisant, UK). Prior to sample collection TD tubes were conditioned at 325°C for 40 minutes in a stream of nitrogen passed through a hydrocarbon trap (Supelco, US) using a Markes International TC-20 tube conditioner (Markes International, Llantrisant, UK). Details of the conditions of analysis using TD-GC-MS have been published elsewhere[84]. Briefly, TD tube samples were pre-purged for 1 min at 50 mL/min constant helium flow rate prior to 280°C for 10 min. Following secondary desorption by heating the cold trap (U-T12ME-2S) from 10 °C to 290°C at 99°C/min and held for 4 min. The GC flow path was heated constantly at 140°C. VOC separation was performed on a ZB-624 capillary column (60 m × 0.25 mm ID × 1.40 µm df; Phenomenex Inc., Torrance, USA) programmed at 1.0 mL/min constant Helium carrier flow. Oven temperature profile was set at 40°C initially for 4 min, ramp to 100°C (5°C/min with 1 min hold), ramp to 110°C (5°C/min with 1 min hold), ramp to 200°C (5°C/min with 1 min hold), finally ramp to 240°C at 10 °C/min with 4 min hold. The MS transfer line was maintained at 240°C whilst 70 eV electron impact at 230°C was set while the quadruple was held at 150°C. MS analyser was set to acquire over the range of 20 to 250 m/z with data acquisition approximated to 6 scan/sec. GC-MS data was then processed using MassHunter software version B.07 SP1 (Agilent Technologies, Cheshire, UK) while MS data of the separated VOC component was compared with NIST Mass Spectral Library version 2.0 for identification [85].

5.3.2.5 *Statistical analysis*

Statistical analysis was performed using Prism (Ver. 7.0d, GraphPad Software, San Diego, CA) and SPSS (Ver. 21, IBM Corp., Armonk, USA). GC-MS signals for detected VOCs (not normally distributed) are presented as medians and interquartile range. Mann-Whitney's U test was performed to compare two independent groups, and Kruskal–Wallis test to compare multiple groups. Statistical significance was assigned to two-tailed *P*-values that were <0.05.

5.3.3 Results

In total 25 patients with oesophago-gastric cancer (17 male, 74±14yrs) and 20 control subjects (10 male, 57±17yrs) were recruited. Baseline sampling of mixed breath using the ReCIVA device was completed in all patients and an additional isolated bronchial breath sample was collected in all cancer patients. In two patients, intraluminal gastric headspace sampling was abandoned due to contamination of the sampling line with gastric secretions. The median peak areas of the different VFAs in these compartments are presented in Table 7. ROC analysis for butyric and pentatonic acid within the endoluminal gastric air, presented in Figure 9, gave an area under the curve of 0.80 (95% CI 0.65 to 0.93; *P*=0.01). Unsupervised PCA and supervised OPLS analysis demonstrated that the examined VFAs contribute to the clustering and discrimination of gastric endoluminal air between cancer and control subjects as shown in Figure 10.

	Cancer mixed breath N=23	Control mixed breath N=20	P^1	Cancer bronchial breath N=25	P^2	Cancer Gastric endoluminal N=23	Control Gastric endoluminal N=20	P^3
Acetic acid	364 (280-567)	125 (7-425)	0.006	304 (208-439)	0.170	750 (319-1020)	747 (297-1076)	0.344
Buytric acid	16 (7-35)	4 (1-24)	0.047	16 (8-24)	0.653	157 (109-192)	49 (20-104)	<0.001
Pentanoic acid	8 (3-11)	4 (0-11)	0.026	8 (4-13)	0.644	144 (36-152)	31 (75-80)	0.005
Hexanoic acid	6 (2-12)	7 (2-30)	0.620	12 (5-17)	0.082	149 (52-217)	64 (24-252)	0.511

Table 7:Median values of peak areas (counts rate $x10^3$) of volatile fatty acids in different aerodigestive compartments between patients and controls. P^1 = mixed breath, cancer vs. control; P^2 = bronchial breath vs. mixed breath in patients with cancer; P^3= endoluminal air, cancer vs. control.

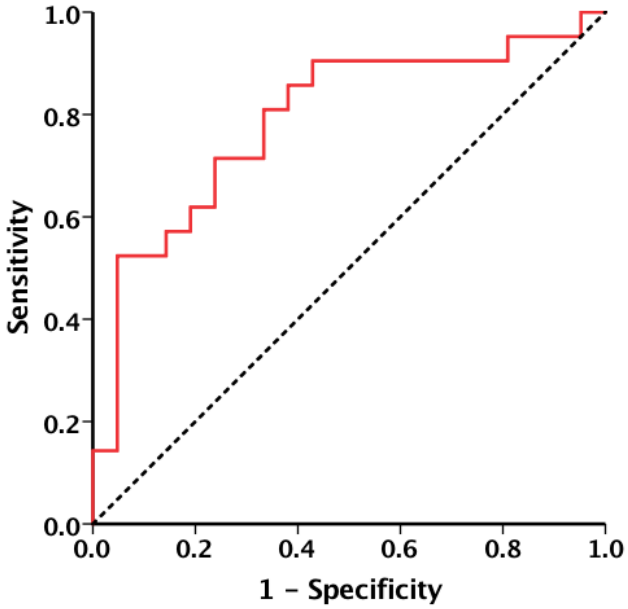

Figure 9:Receiver operating characteristic curve for gastric endoluminal volatile fatty acids significant on univariate analysis; butyric acid and pentatonic acid area under the curve of 0.80 (95% CI 0.65 to 0.93; *P*=0.01).

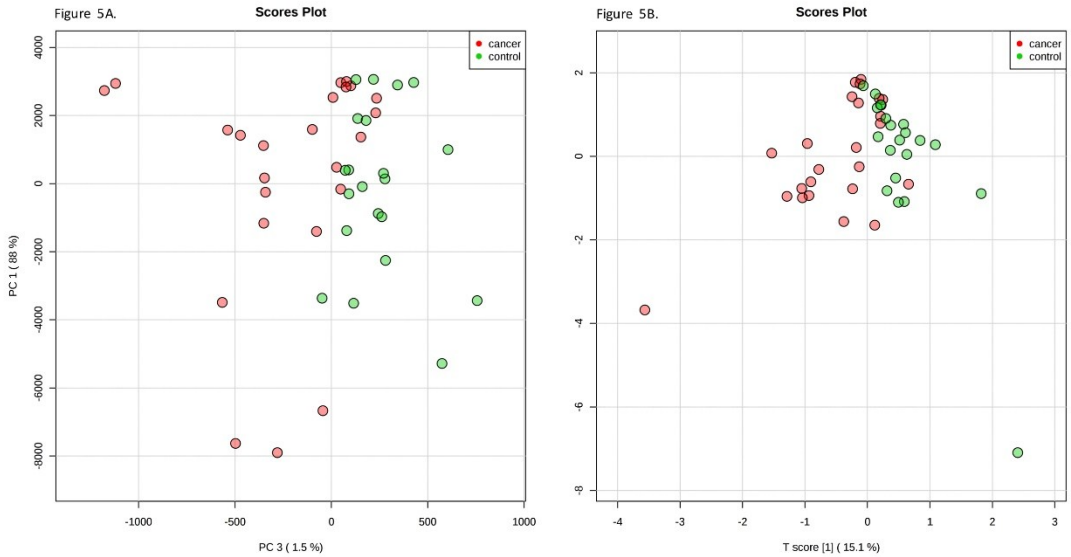

Figure 10:Principal Component Analysis (A) and Orthogonal Partial Least Square (B) of gastric endoluminal volatile fatty acids in cancer and control patients.

Compared to mixed and tracheal breath samples, all examined fatty acids were found at highest concentrations within the gastric headspace. Fatty acids tended to be higher in all samples derived from cancer patients compared to controls.

Butyric acid and pentanoic acid were found to be significantly elevated in the mixed breath and gastric endoluminal air of cancer patients compared to controls, with endoluminal levels being approximately ten times greater than found in mixed breath. Equivalence of VFA levels within mixed and bronchial breath samples from cancer patients suggests that their origin within breath is principally derived from the lungs and by inference the systemic circulation as opposed to direct passage from the upper gastrointestinal tract.

In cancer patients, levels of butanoic acid were significantly higher within luminal gastric headspace compared to isolated tracheal air ($P<0.001$) and mixed breath $P<0.001$). Compared to cancer patients the levels of butanoic acid were significantly lower in both the gastric headspace ($P<0.001$) and whole breath ($P=0.026$) of controls. Details are shown below in Figure 11.

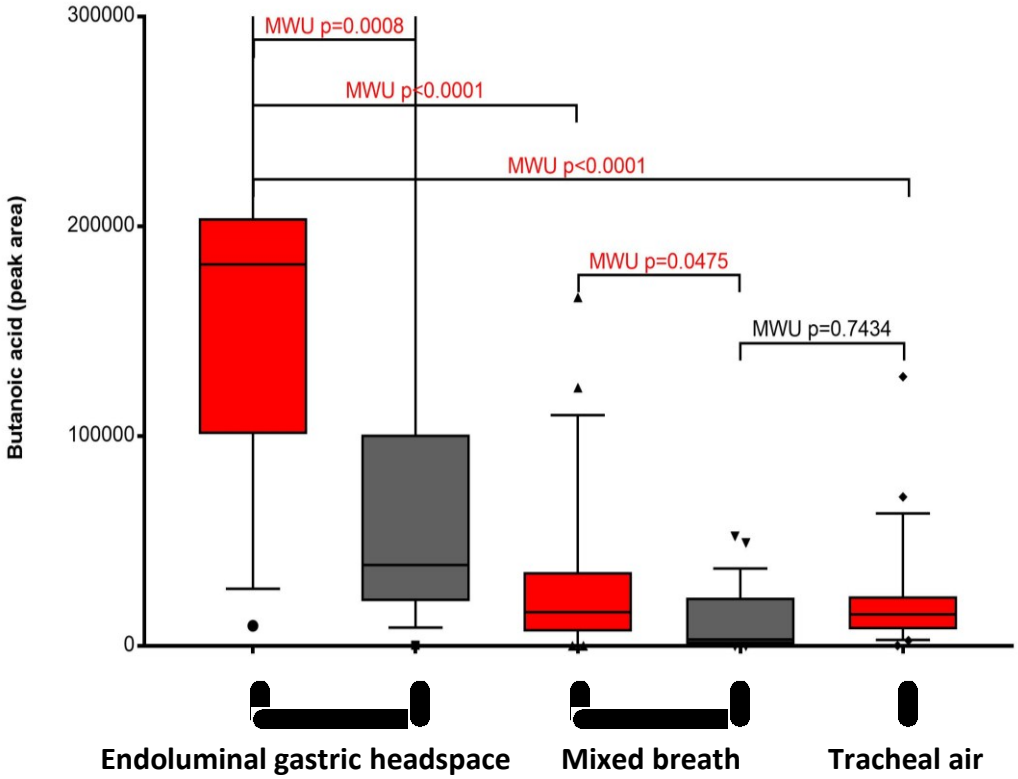

Figure 11:Box plots depicting the concentrations of butanoic acid within different gas compartments of oesophago-gastric cancer (red) and control (grey) subject. Pairwise comparison of groups performed using Mann Whitney U (MWU) test.

A similar pattern was observed for pentanoic acid, which was highest in the luminal gastric headspace of cancer patients compared to isolated tracheal (P<0.001) and mixed breath (P<0.001) (Figure 4). Pentanoic acid levels were significantly lower in the gastric headspace (P=0.013) and whole breath (P=0.048) of controls compared to cancer patients. Details are shown below in the Figure 12.

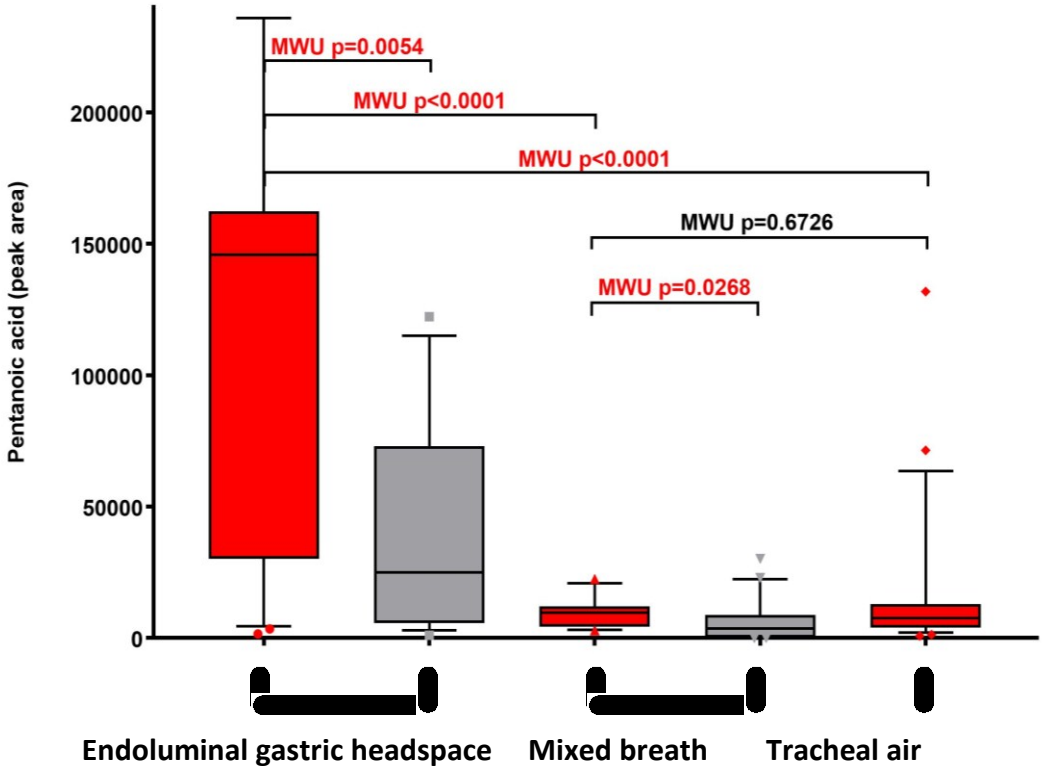

Figure 12:Box plots depicting the concentrations of pentanoic acid within different gas compartments of oesophago-gastric cancer (red) and control (grey) subject. Pairwise comparison of groups performed using Mann Whitney U (MWU) test.

It is noteworthy that whilst acetic acid levels were significantly elevated in the mixed breath of cancer patients, equivalent enriched levels were found in gastric endoluminal air of both cancer and control subjects. This could suggest that the raised levels of acetic acid found within the exhaled breath of patients with oesophago-gastric cancer may be influenced by other, as yet undetermined, systemic sources.

5.3.4 Discussion

This is the first study to perform isolated gas sampling of gastric luminal headspace and tracheal breath in order to investigate the origins of cancer specific compounds. Previous studies suggested multiple possible origins of butanoic and pentanoic acids with whole breath including cancer and colonic bacteria[86,87,21]. Findings presented herein would appear to support that the stomach is the source elevated levels of selected VOCs that have been previously been associated with oesophago-gastric cancer. It is compelling that these compounds are found in higher concentrations within the exhaled breath of cancer patients, but at much lower levels.

As previously mentioned, there are presumed to be two main pathways by which may be excreted to oral breath, either through passage from the systemic circulation across the alveolar capillary barrier or through direct release from the upper airways and digestive tract[59]. Importantly, this study was able to measure isolated tracheal breath in intubated cancer patients. Whilst acknowledging the inconsistency in methodologies used to assess breath in intubated and spontaneously breathing patients the broadly equivalent, but low levels, of cancer markers in these samples may suggest two things. Firstly, whilst these VOCs may be found in relative abundance within the stomach lumen, this does not appear to be a source of significant contamination of exhaled breath. Secondly, if the tumour is indeed the source of these VOCs in exhaled breath the process whereby they are transported to the lung within the systemic circulation before being partitioned across the alveolar capillary barrier leads to a significant attenuation in their detectable levels.

These factors highlight one of the major challenges associated with breath analysis, which is the extrapolation of breath data to extra-pulmonary events. Whilst organ specific sampling at the site of disease reduces the potential for both the dilution and contamination of target VOCs, in many cases this approach may be impractical and negate the principal benefit of breath analysis, which is its status as a non-invasive test. The focus should therefore remain on the optimisation of breath testing in areas where there is a defined clinical need, such as the early detection of oesophago-gastric cancer. That having been said, there may be an as yet unknown future role for the rapid confirmation of gastrointestinal disease at the time of endoscopy by way of a 'gas biopsy'.

The absorption of fatty acids occurs within the small and large bowel in humans. They have multiple roles following absorption including important immune modulation, minor nutritional source, altering pancreatic islet cell function and many other processes that are not fully understood yet[88]. Changes to fatty acid synthesis resulting apoptosis in different type of cancers including gastric malignancy has been widely investigated before[89]. The most likely mechanism relies in the over-expression of the fatty acid synthase enzyme and the consequent de novo synthesis of fatty acids[90,91]. Over expression of this enzyme also linked to poor outcome from the disease[92].

The current study has contributed new evidence regarding the stomach as the sources of aberrant fatty acids levels in oesophago-gastric cancer, however it does not precisely define where these VOCs are emanating from. It is noteworthy that whilst fatty acids levels were seen in highest concentrations in the luminal headspace of patients with oesophago-gastric cancer, they were readily detectable within the stomach of healthy subjects. Therefore,

production is not only cancer specific and may reflect either the deregulation of normal physiological production within gastric cells and/or the gut microbiota[93].

This is the first study to have attempted to analyse VOC levels from three defined body compartments: whole breath, tracheal breath and luminal headspace. Whilst it was intended that a 500mL gas sample were taken from each compartment variation in the conditions of sampling mean that caution should be taken when trying to infer direct correlations between the relative abundance of VOCs. Further studies are needed to refine the methodology for organ specific detection of VOCs and to define the kinetics of their release into the gas phase. Culturing experiments would also allow for a better understanding of the source of VOC production at a microbial level.

5.3.5 Conclusion

There are diagnostic clinical implications of these studies. The marked difference in VFA levels in the gastric endoluminal air of cancer compared to control patients offers the opportunity for an endoluminal gas biopsy for cancer detection. Secondly, the non-significant difference between exhaled and isolated bronchial breath supports the use of mixed exhaled breath for non-invasive cancer detection without the need for complex devices for alveolar sampling. In conclusion, this study has contributed to existing evidence that fatty acids are deregulated in oesophago-gastric cancer. Whilst results would appear to confirm that the stomach, and presumably the tumour, is the source of these VOCs further work is needed to define the mechanisms of their production and release.

6 ASSESSMENT OF MICROBIAL DIVERSITY IN OESOPHAGO-GASTRIC CANCER

6.1 Background

There has been growing interest in the role of microbiome and its contribution to human disease. Concurrent advances in both the methods by which components of the microbiome are determined and defined has help to support this growing field of research. The terms microbiome and microbiota reflect a community of commensal, symbiotic and/or pathogenic microorganisms within multicellular host organism. The microbiome describes all bacteria, fungi and viruses and their collective genomes found within the host. Studies have shown that the microbiome has an important role in regulating host immune, hormonal and metabolic function [94].

Traditionally the stomach was considered to be a sterile organ on account of its secretion of hydrochloric acid proteolytic enzymes. There is now however, growing evidence for the role of the upper gastrointestinal microbiome in the development of oesophago-gastric cancer. This association is thought to reflect the microbiomes effect on inducing and maintaining the carcinogenic pathways by stimulation of inflammation, a rise in cell proliferation and the dysregulation of stem cell physiology[95].

Helicobacter pylori (*H pylori*) is the most well-known and studied gastric bacteria. Research about *H pylori* has advanced the understanding of how bacteria can modify its own microenvironment and induce a disease state[96].

Non-*H pylori* bacteria may also play a role in the development of gastric cancer. Studies showed that in male mice with human intestinal microbiota developed chronic gastritis with subsequent progression to gastric atrophy and dysplasia that was independent of *H pylori*

infection. The presence of complex commensal microbiota accelerated the progression to gastric intraepithelial neoplasia in *H pylori*–infected mice compared to mice only infected by *H pylori* [97,98].

Further investigations with prevailing techniques have shown that the microbiota of the stomach involves hundreds of phylotypes with a microbial concentration between 10^1 and 10^3 colony-forming units (CFU)/g [94,104]. However, like the rest of the intestinal microbiome, this is a dynamic and complex system with substantial variations in microbial density changing with pH, whereby both the number and quantity of the genera fluctuate. Human gastric content has an interprandial luminal pH of between 1 and 2, although may increase above 5 following food ingestion. Such fluctuations in gastric pH may drive changes in the microbiome.

Few studies have examined the role of the microbiome in oesophageal and gastric disease. Gouquin *et al.* have demonstrated that individuals with lower microbial diversity were prone to have chronic atrophic gastritis and squamous dysplasia in the oesophagus [99]. Another study using high-throughput sequencing methods (454 GS FLX Titanium) demonstrated a relative increase of Baccili and the streptococci species in the cancer group with a relative reduction of *H pylori* [50].

Whilst a number of studies have reported that the microbiome in individuals with oesophago-gastric cancer and precancerous conditions is altered [100], there is relatively little evidence as to which species of bacteria are altered in these states and whether chemotherapy has an impact on microbial changes. Chemotherapy is widely used in the treatment of oesophago-gastric cancer. Emerging evidence suggests that the gut microbiota

may play a role in this response. Drug metabolism by the intestinal microbiota has been recognised in the past as playing an important role in influencing the drug efficacy and toxicity[101]. The Intestinal microbiota was found to influenced fluropyrimidines, which are commonly used in the treatment of colorectal cancers[102].

Work described within this chapter intends to better define the specific bacterial species associated with oesophago-gastric cancer and the potential effects of chemotherapy on the gastric microbial community.

Hypothesis:

It is hypothesised that oesophago-gastric cancer will be associated with a unique tumour associated microbiome, distinguishing it from both surrounding healthy mucosa and mucosa from healthy controls.

Aims:

The primary aim of this study is to characterise the microbial composition of patients with oesophago-gastric cancer.

Secondary aims are to:

i. To define and confirm the presence of the most abundant bacteria present within oesophago-gastric adenocarcinoma tissue.

ii. Determine the effect of chemotherapy on the gastric microbial structure.

iii. To detect the presence of specific genes responsible for producing target VOCs found to be upregulated in oesophago-gastric cancer patients.

6.2 **Initial pilot study**

Before conducting a more definitive assessment of the upper gastrointestinal microbiome of oesophago-gastric cancer and control patients, an initial pilot study was performed to establish specific protocols and optimise parameters needed to perform the principal study.

6.2.1 Aims

i. To establish a reliable protocol for the collection and processing of oesophago-gastric specimens intended for microbiome analysis
ii. To identify the abundance of bacterial species within gastric adenocarcinoma tissue compared to adjacent healthy mucosa taken from the same patient.

6.2.2 Methods

Subjects:

Ten subjects with biopsy confirmed oesophago-gastric adenocarcinoma were recruited to this pilot study. The exclusion criteria were patients with known liver disease, existing small bowel/colonic conditions or any other form of cancer other than oesophago-gastric adenocarcinoma. Patients were requested to provide a comprehensive medical history, smoking status and alcohol intake history. Ethical approval for this study was granted by Imperial College London (Ref. R16035-6A).

Sample collection:

Specimens were obtained at the time of elective upper gastrointestinal endoscopy or surgical resection. Samples were taken from the tumour as well as adjacent macroscopically normal stomach mucosa by either cold biopsy forceps (2.8mm EndoJaw, Olympus, UK) or direct excision using a sterile scalpel blade. Samples were collected during routine endoscopy or at the time of oesophago-gastric cancer resection.

Following collection tissues samples were divided equally with one portion placed in formalin for histological assessment and the remaining tissue being snap frozen in liquid nitrogen before being stored at -80^{o}C. Details of sample characteristics are presented in in Table 8.

Subject No.	Method of sample collection	Tumour Site	Tumour Stage	Biopsy type	Biopsy Weight (mg)	Qubit (ng/ul)	Neo-adjuvant therapy	Mandard score
1	TG	Cardia	T1N1MX	NS	84	1.7	Yes	5
				C	28	2.3		
2	TG	Lesser curve	T0N0MX	NS	55	1.0	Yes	1
				C	13	0.9		
3	Endoscopy	Lesser curve	T3N0MX	NS	5	0.4	No	N/A
				C	5	0.5		
4	TG	Body	T3N1MX	NS	50	0.8	Yes	2
				C	153	1.7		
5	PG	Antrum	T3N0MX	NS	26	1.6	No	N/A
				C	31	0.5		
6	TG	Lesser curve	T4N2M1	NS	39	0.9	Yes	4
				C	51	4.9		
7	SL/Endoscopy	Cardia	T4N2M1	NS	152	2.4	Yes	4
				C	23	0.9		
8	TG	Body	T3N1M0	NS	19	1.4	Yes	3
				C	23	2.7		
9	SL/Endoscopy	Lesser curve	T4N3MX	NS	8	3.8	Yes	3
				C	3	2.2		
10	TG	Lesser curve	T3N0MX	NS	18	1.1	Yes	2
				C	22	0.8		

Table 8:Characteristics of tissue samples and amount of DNA extracted (Qubit). TG, total gastrectomy. PG, partial gastrectomy. SL, staging laparoscopy. NS, normal stomach. C, cancer. Qubit, Amount of DNA extracted prior to sequencing.

6.2.2.1 *DNA extraction*

To extract genomic DNA, biopsies were first thawed before being homogenised and treated with 1U of DNase I to eliminate any potentially foreign bacterial DNA. Genomic DNA was extracted using the MOBIO DNA Micro Kit (MoBio, Carlsbad, CA, USA) with concentration measured using Qubit 2.0 fluorometer. Tissue samples were added to glass beaded tubes, where homogenization and lysis took place by adding different beading solutions aiming to

breakdown fatty acids and lipids associated with cell membrane of several organisms. Homogenisation was performed by gently using the vortex followed by centrifugation at 14,000xg for 10 minutes to pellet cell debris. Supernatant was then suspended in a tube with reagents to precipitate additional non- DNA organic and inorganic material including proteins. Centrifugation performed at 140,000 xg for 1 minute was performed to pellet cell debris. This process was repeated again with ethanol to precipitate additional non-DNA products and incubated at 4^{o}C for 5 minutes. 1200 µl of high concentration salt solution was added to supernatant and the sample mixed for 1 minute. The mixture was transferred to a SPIN^{TM} Filter and centrifuged at 10,000xg for 1 minute. To purify the sample, the process was repeated three times. The SPIN^{TM} filter was then air-dried at room temperature before addition of elution buffer which results in a more efficient and complete release of DNA from the silica spin filter membrane. Centrifugation at 10,000 xg for 1 minute brought the eluted DNA into a clean tube and sample was stored at -80^{o}C until subsequent analysis.

6.2.2.2 *Real-time Polymerase Chain Reaction (PCR)*

To determine the bacterial load and gene abundance in tissue mucosa, qPCR is considered to be the most accurate and effective method of library quantitation, providing considerably high consistency and reproducibility. Quantification of the samples was made using the New England BioLabs library quantification kit qPCR for Illumina utilizing Sybr green reagents. Preparing the library using the steps to amplify the V3 and V4 region and using a limited cycle PCR, up to 96 libraries was pooled together for sequencing. Two stages of PCR were used and in between there are clean-up methods. GeneAmp PCR system 9700 cycler was used with cycling parameters of:

i. Initiation (95^{o}C for 3 minutes)
ii. Denaturation (25 cycles at 95^{o}C for 30 seconds) to separate the DNA strands
iii. Annealing (25 cycles at 55^{o}C for 30 seconds) to allow primers to base pair to complementary DNA template
iv. Elongation (25 cycles at 72^{o}C for 30 seconds)
v. Final extension (72^{o}C for 5 minutes)

Amplicon PCR (**First stage**):

25 µl PCR reactions were set up with 2.5 µl of microbial DNA, 5 µl of Amplicon Forward Primer (1µM), 5 µl of Amplicon Reverse primer (1 µM), 12.5 µl 2x KAPA HiFi HotStart Ready Mix).

Index PCR (**Second stage**):

50 µl PCR reactions were set up with 5 µl of Nextera XT Index Primer 1 (N7xx), 5 µl of Nextera XT Index Primer 2 (S5xx), 25 µl of 2x KAPA HiFi HotStart Ready Mix and10 µl PCR Grade water.

6.2.2.3 *16S rRNA Sequencing*

Metagenomic studies are commonly performed by analysing the prokaryotic 16S ribosomal RNA gene (16S rRNA), which is approximately 1,500 base pairs long and contains nine variable regions interspersed between conserved regions. Variable regions of 16S rRNA are frequently used in phylogenetic classifications such as genus or species in diverse microbial populations. The most appropriate 16S rRNA region to sequence is an area of debate, and the region of interest might vary depending on factors including experimental objectives, design, and sample type.

This pilot study used the Illumina (Illumina Inc., Saffron Walden, UK) protocol, which sequences the variable V3 and V4 regions of the 16S rRNA gene. The protocol included the amplicon primer pair sequences specifically for V3 and V4 region that creates a single amplicon of approximately 460 base pairs. The protocol also included overhang adapter sequences and dual-index barcodes that must be appended to the primer pair sequences for compatibility with Illumina index and sequencing adapters.

Library normalization was done using sequel prep normalization plate (96) kit eliminating the amplicon quantitation and manual normalization steps. It allowed a single step, high-throughput amplicon purification and normalization of PCR product concentration (2-3 fold range) through a limited binding capacity solid phase. PCR products were added to a SequelPrep plate well, each well can bind and elute 25ng of PCR amplicon and then mixed with the binding buffer. DNA binding to the plate is performed at room temperature for one hour. The wells were washed with wash buffer to efficiently remove contaminants. Purified PCR products were eluted using 20 µl elution buffer at normalized concentrations. Samples after normalization were pooled using equimolar amounts of all DNA from each sample. After knowing the library concentration, samples are diluted down to the concentration desired for the Miseq loading.

Samples were subjected to metataxonomic analysis on the Illumina MiSeq platform, with the V3/V4 region of oesophago-gastric cancer microbiomes being targeted in a high-multiplexing approach, thus leading to a high coverage of the microbial diversity. Sequencing on MiSeq, using paired 300-bp reads, and MiSeq v3 reagents, the ends of each read are overlapped to generate high-quality, full-length reads of the V3 and V4 region in a

single 65-hour run. The MiSeq run output is >20 million reads and commonly recognized as sufficient for metagenomic surveys.

6.2.2.4 *Analysis*

Taxonomic-dependent analysis of reads from amplicon sequencing was performed using *Mothur software* following the MiSeq SOP Pipeline. The non-metric multidimensional scaling (NMDS) plot and PERMANOVA p-values were generated using the UniFrac weighted distance matrix generated from Mothur, and analysed using the Vegan library within the R statistical package. Statistical tests were performed using IBM SPSS Statistics Software version 23. A p value of <0.05 was considered significant.

6.2.3 Results

The results demonstrated consistency with previous studies showing the dominant bacterial phyla within the stomach were Firmicutes and Bacteroidetes (Figure 13,14). The total amount of DNA extracted from tissue samples was extremely low, making interpretation difficult and limiting the strength of conclusions. A batch effect was observed in one set of samples showing a large variation in the number of operational taxonomy units within each sample.

The approach adopted did not enable a definitive conclusion to be made and did not assist in determining whether there were any differences in relative abundance between cancer and normal stomach from the same cancer patient. These results however provided greater insight into the methodological challenges of microbiome analysis that were used to optimise protocols for the principal study.

Key learning points for optimisation:

- Low bacterial DNA concentration present in biopsy samples
- Eliminate the possibility of batch effects
- Effect of chemotherapy

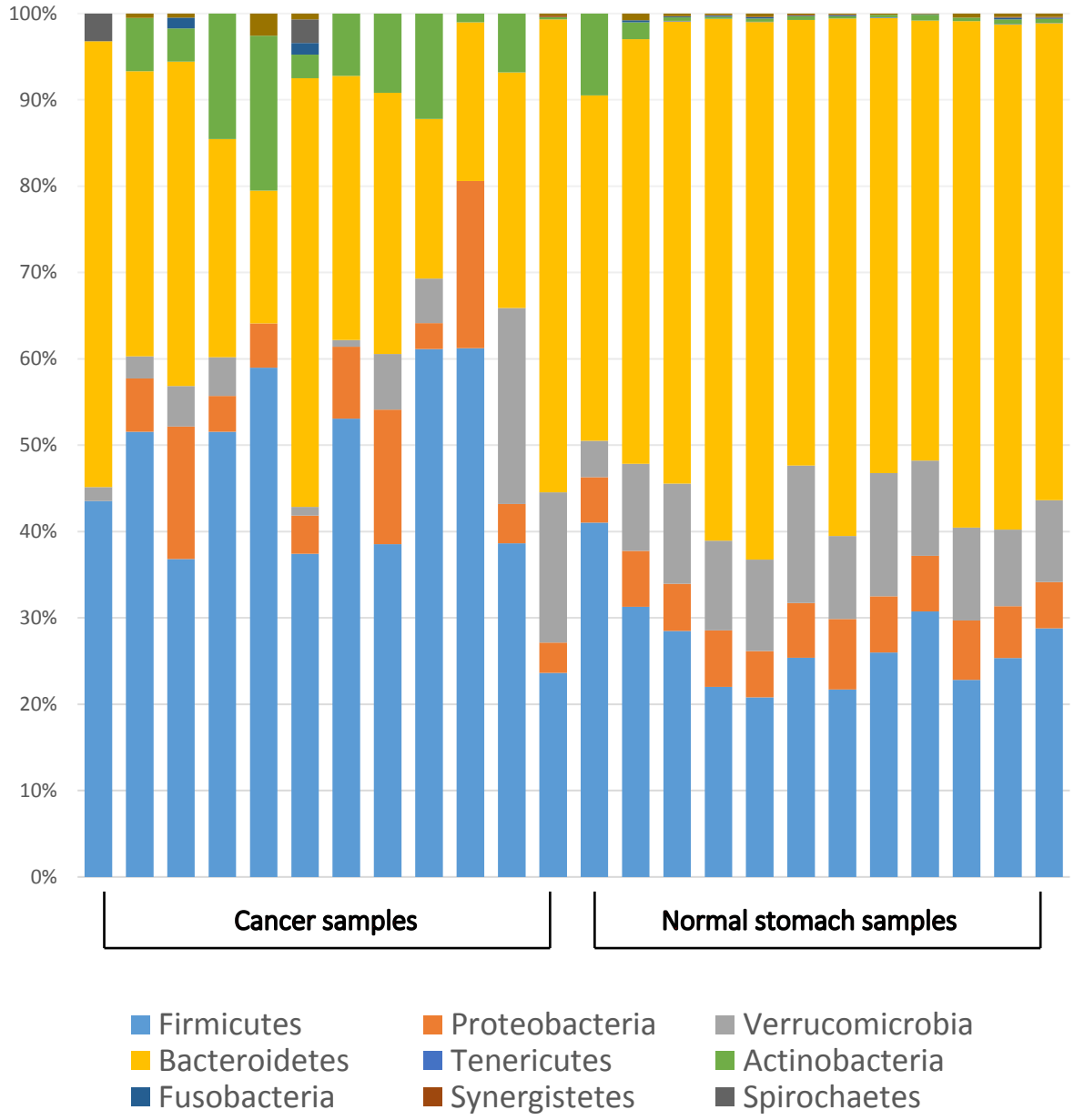

Figure 13:Differences in phylum level between cancer and normal stomach samples from the same cancer patient. Firmicutes and Bacterioidetes are mostly present in all samples.

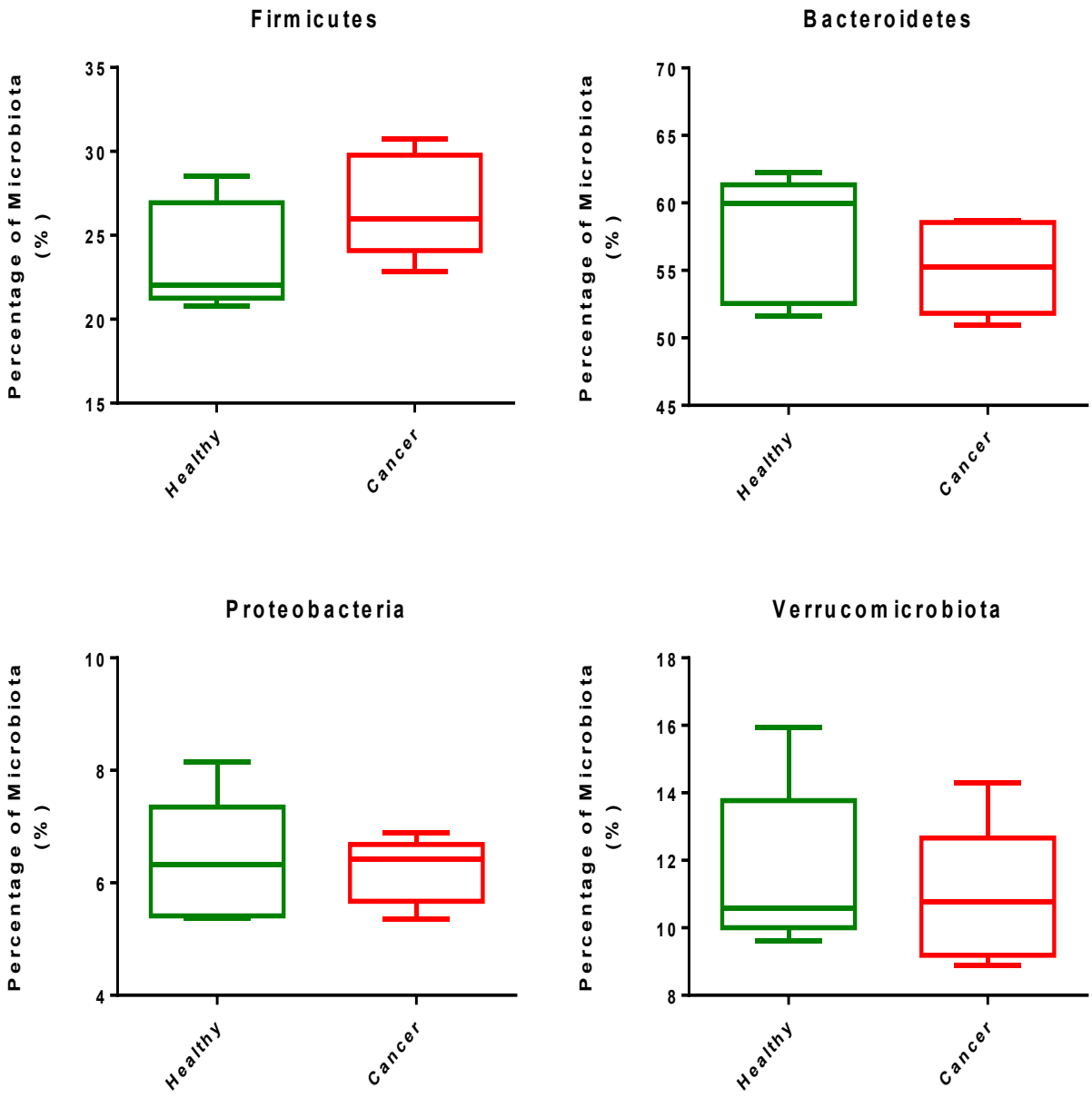

Figure 14:Box plots comparing the phylum level of bacteria present in cancer and control patients. Firmicutes (p=0.075), Proteobacteria (p=0.787), Bacteroidetes (p=0.126), Verrucomicrobiota (p=0.169).

6.2.4 Discussion

The microbiome of the stomach and the oesophagus is currently poorly understood. Gastric content is mainly composed of proteolytic enzymes and hydrochloric acid, which restricts the quantity of microorganisms entering the small intestine and reduces the risk of infection by pathogens [103]. In certain disease states, reduction of gastric acid may influences the composition of intestinal or oral microorganisms, including pathogenic organisms [104]. Little is currently known regarding the microbial composition of the stomach of patients with gastric cancer. Furthermore, there remains uncertainty as to the optimal methodology for retrieval and analysis of oesophago-gastric bacteria. A major objective of this this pilot study was to determine the efficacy of bacterial retrieval under standard methods and to assess whether amplification was required in order to establish a robust method for future studies.

In this pilot study, the oesophago-gastric cancer microbiome was characterised using genomic sequencing approaches and compared with normal matched samples from the same cancer patients. Representative bacteria of several genera were detected with a predominance of Firmicutes in cancer samples. The primary outcome demonstrated several other bacteria (Bacteroidetes, Proteobacteria and Verrucomicrobiota) to differentiate between cancer and normal stomach samples, however there were no significant differences in microbial composition between both.

To address the low bacterial DNA concentration present in samples, several techniques have been introduced in the method to amplify the DNA during extraction. This includes increased concentrations of reagents in particular using 50µl of sterile elution buffer instead of 25µl forming the final stage of extraction. This will result in a more efficient and complete release of DNA with higher concentrations. Also to eliminate the possibility of batch effects

in the main study, the Bullet Blender Storm (Chembio Ltd, St Albans,UK) was introduced to exclude the need for manually manipulating samples between different cylces using the vortex machine. The Bullet Blender can homogenise samples equally at high acceleration in only 45 seconds using the glass beads to achieve cell lysis efficiently in less time.

This study intended to look at gastric cancer only without healthy controls. Investigation of the effect of chemotherapy, and other factors, on the microbiota was not an intended outcome of this study and will be addressed in the principal study. Although, there were a limited number of subjects examined, this study contributed to the principal study, by exploring the complex bacterial community composition associated with oesophago-gastric cancer patients.

6.2.5 Conclusion

The main aim for this exploratory study was to assess the microbial diversity and to establish the appropriate method for 16S sequencing on human tissue samples. Although no significant differences between proportions of specific bacteria were demonstrated in this pilot study, but the method helped in determining differences in relative abundance of some phyla and to enhance the approach for a larger study. The numbers of patient samples included were small, and the larger planned study described in the next section was intended to provide more definitive conclusions.

6.3 **Principal study of the microbial profile of oesophago-gastric cancer**

Characterisation of the tumour associated microbiome is intended to provided further evidence for the role of bacteria in the pathogenesis of esophago-gastric cancers. In addition those bacteria that are disproportionally associated with esophagogastric cancers are hypothesised to be the source of elevated VOCs that are detected within the exhaled breath of this patient group [19-22].

6.3.1 Aims

i. Using an optimised methodology to confirm the abundance of bacterial species within esophagogastric adenocarcinoma tissue.
ii. To compare the microbial profile of esophagogastric adenocarcinoma tissue to that of adjacent healthy gastric and oesophageal mucosa as well as mucosa from healthy controls.
iii. To determine the effect of chemotherapy on microbial diversity.

6.3.2 Methods

6.3.2.1 *Study participants*

Tissue samples were obtained during planned upper gastrointestinal endoscopy. In oesophago-gastric cancer patients this was performed at the time of staging laparoscopy

prior to patients commencing neo-adjuvant chemotherapy or any other form of treatment for their underlying cancer. Tissue biopsies were acquired using cold biopsy forceps (2.8mm EndoJaw, Olympus, UK) from the central portion of the tumour. In oesophageal cancer patients biopsies were also acquired from macroscopically normal stomach and oesophageal mucosa, however in gastric cancer only normal stomach mucosal samples were taken. For control patients, biopsies were acquired at the time of routine upper gastrointestinal endoscopy, from the fundus in control patients using cold biopsy forceps. Biopsies were immediately snapped frozen in liquid nitrogen prior to being transferred to a -80^{o}C freezer for storage prior to analysis.

To study the effect of chemotherapy, tumour samples were collected from 14 gastric cancer patients at the time of resection of the specimen. Following resection of the stomach the specimen was opened and biopsies were acquired using a sterile scalpel blade. These patients had paired biopsies collected at the time of staging laparoscopy before neo-adjuvant chemotherapy.

Local ethics committee approval was granted for this study (Ref:12/WA/0196) and written informed consent was obtained from all patients prior to enrolment in the study.

6.3.2.2 *DNA extraction and 16S rRNA gene sequencing*

DNA was extracted from tissue samples using the PowerLyzer PowerSoil DNA Isolation Kit (MoBio, Carlsbad, CA, USA) following manufacturer's instructions, with the addition of the homogenising step using Bullet Blender Storm (Chembio Ltd, St Albans,UK) for 3 min at speed 8 in a DNA was stored at −80 °C until it was ready to be used.

Sample libraries were prepared following Illumina's 16S Metagenomic Sequencing Library Preparation Protocol with two modifications. Firstly, the V3-V4 regions of the 16S rRNA gene were amplified using the primers listed in Table 9. Additionally, the index PCR reactions were cleaned up and normalised using the SequalPrep Normalization Plate Kit (Life Technologies, Paisley, UK). Sample libraries were quantified using the NEBNext Library Quant Kit for Illumina (New England Biolabs, Hitchin, UK). Sequencing was performed on an Illumina MiSeq platform (Illumina Inc., Saffron Walden, UK) using the MiSeq Reagent Kit v3 (Illumina) and paired-end 300 bp chemistry.

Primer name	**Primer sequence**
28F-YM (forward primer)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAGTTTGATYMTGGCTCAG
28F-Borrellia (forward primer)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAGTTTGATCCTGGCTTAG
28F-Chloroflex (forward primer)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGAATTTGATCTTGGTTCAG
28F-Bifdo (forward primer)	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGGTTCGATTCTGGCTCAG
388R	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGCTGCCTCCCGTAGGAGT

Table 9:Primers used for 16S rRNA gene sequencing on the Illumina MiSeq. The forward primer mix was composed of four different forward primers, mixed at a ratio of 4:1:1:1 (28F-YM:28F-Borrellia:28FChloroflex:28F-Bifdo). Bases in bold are the MiSeq adapter sequences.

6.3.2.3 *Real-time PCR*

qPCR was performed using extracted DNA to quantify gene abundance. Gene abundance was quantified for specified groups using degenerate primer sets designed and optimised as seen in table 10. A total reaction volume of 25 µl was used for each reaction, consisting of 20 µl master mix and 5 µl diluted DNA (12.5 ng total per reaction). All DNA was diluted in buffer EB (Qiagen, Hilden, Germany). A standard master mix consisting of 5.5 µl PCR grade

water (Roche, Penzberg, Germany), 12.5 µl of 2x SYBR green master mix (ThermoFisher Scientific, Waltham, Massachusetts, USA), 1 µl of 10 µM forward primer (Eurofins Genomics, Wolverhampton, UK) and 1 µl of 10 µM reverse primer (Eurofins Genomics) was used. One bacterial strain from the relevant reference group was selected as a standard for each primer set. Serial dilutions of each isolate were used to create a standard curve. Thermocycling conditions for each primer set are summarised in Table 10.

A melt curve stage was performed post-cycling to confirm primer specificity. Products were also visualised using the 2200 Tapestation System (Aligent Technologies, Santa Clara, California, USA) in combination with D1000 Reagents and D100 Screentapes (Aligent Technlogies), following the manufacturer's protocol.

Copy number was calculated from qPCR data using the following formula: gene abundance=(quantity (ng) x6.022 $x10^{23}$ (gene copy number/mol))/(length of product x 1x 10^{9} (ng/g) x 660 (g/mol)). A mean copy number for each set of triplicates was calculated and divided by the total DNA per reaction to obtain average copy number per ng DNA.

GROUP	Primer sequence (5'-3')	F/R	Cycling conditions	Expected product size (bp)
1a	CACATATTGTGGCACGA ACAATHGAR TGGGG CTGTGCCCGGATACAGA TTAACRTAR TTRTT	F R	95 °C for 10 min, (95 °C for 15 s, 55 °C for 1 min)Å~40 cycles	570
1b	CGGCGTTCCGCATTTYT AYGARAA GTTCAATGCCAATCGGA ATATCRAAR TTRTT	F R	95 °C for 10 min, (95 °C for 15 s, 55 °C for 1 min)Å~40 cycles	318
2	TTTTGGCCGAACACTGG AYTAYGARTT TCAACGGAGCCCAGAAT ATGRAARA AYTG	F R	95 °C for 5 min, (95 °C for 15 s, 54 °C for 30 s, 72 for 10 min)Å~40 cycles	774
3	GGWTTCAGCCCRCAGA TGTTCTTTG GAATTCCGGGTTCATGA ACATTCTKCKAAG	F R	94 °C for 2 min, (94 °C for 20 s, 52 °C for 30 s, 69 °C for 90 s)Å~35 cycles, 68 °C for 10mins	1300

Table 10**:**Primer sequence and PCR conditions.

6.3.2.4 *Statistical analysis*

The resulting data was analysed using the Mothur package following the MiSeq SOP Pipeline. The Silva bacterial database was used for sequence alignments (www.arb-silva.de/) and the RDP database reference sequence files were used for classification of sequences using the Wang method. The non-metric multidimensional scaling (NMDS) plot and PERMANOVA p-values were generated using the UniFrac weighted distance matrix generated from Mothur, and analysed using the Vegan library within the R statistical package. Family-level extended error bar plots were generated using the Statistical Analysis of Metagenomic Profiles software package using White's non-parametric t-test with Benjamini-Hochberg FDR. The α diversity (Shannon diversity index, H') and richness (total number of bacterial taxa observed, Sobs) were calculated within Mothur and statistical tests (independent t-test and Mann-Whitney U test, respectively) were performed using IBM

SPSS Statistics Software version 23. A p value of 0.05 and a q-value of 0.05 was considered significant.

6.3.3 Results

6.3.3.1 Results of pre-chemotherapy samples

To investigate the microbial composition of gastro-oesophageal adenocarcinoma, 16S r RNA gene amplicon sequencing was employed. In total 50 cancer patients (20 gastric and 30 oesophageal) and 80 healthy controls (50 gastric and 30 oesophageal) were recruited to this study. Details for each group are shown below in Figure 15. The microbial communities of oesophago-gastric mucosa from patients were analysed after filtering the sequencing data and removing contaminant sequences showing alteration of composition associated with cancer tissue compared with matched adjacent normal mucosa from the same cancer patient and healthy controls.

Patients with gastric and oesophageal cancer were older (p=0.001) and predominantly male (p=0.025) compared to control patients. Acid-suppressant drugs was taken regularly by 46 (92%) cancer patients and 52 (61%) control patients (p=0.082). Of the cancer patients the majority of gastric cancer cases had advanced disease with 14/20 having T-stage T3/T4 disease, however the oesophageal samples 16/30 had a staging of T3/T4 and the remaining were T2/T3.

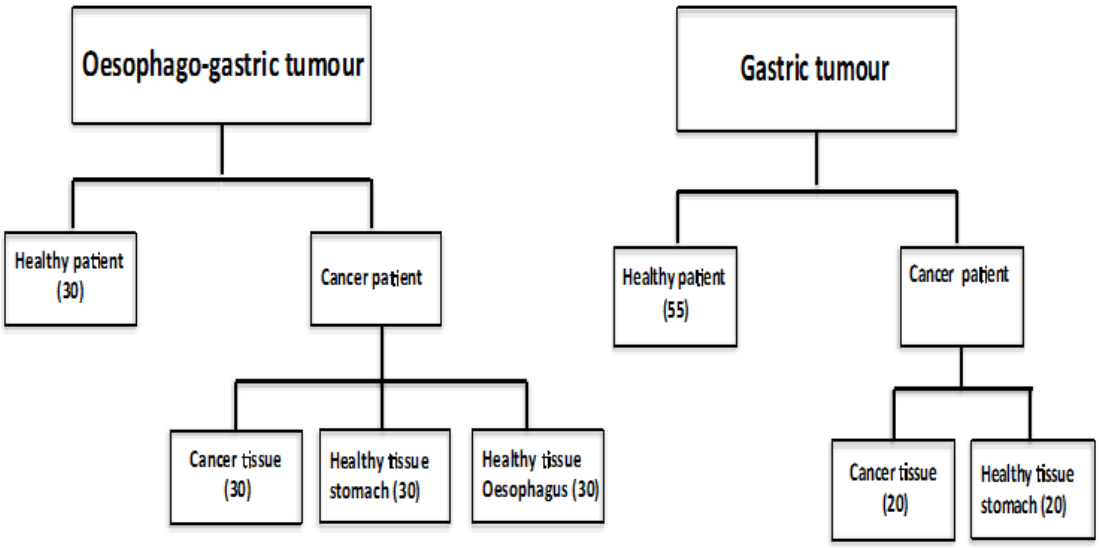

Figure 15:Classification of samples according to their site. Junctional oesophago-gastric tumour samples contained cancer, healthy matched stomach and oesophageal tissue samples. As for Gastric tumour it contained cancer and healthy matched stomach samples only.

6.3.3.1.1 Gastric cancer 16S rRNA sequencing analysis

The microbial communities of gastric mucosa from 75 patients (55 controls and 20 cancers) were analysed using high-throughput sequencing 16s rRNA gene. The Chao1 estimator which is a non-parameteric estimator of species richness, estimated the richness in gastric cancer (283.2) to be significantly lower than that in unmatched healthy controls (309.3, $p < 0.05$). Moreover, there was a similar trend when comparing tumour samples with matched normal gastric mucosal samples from the same cancer patient (318.3, $p<0.05$). However with Inverse simpson index used to quantify the average proportion abundance of types, the diversity of bacteria was shown to be significantly lower when comparing tumour and healthy control patients only, see Figures 16 and 17 below.

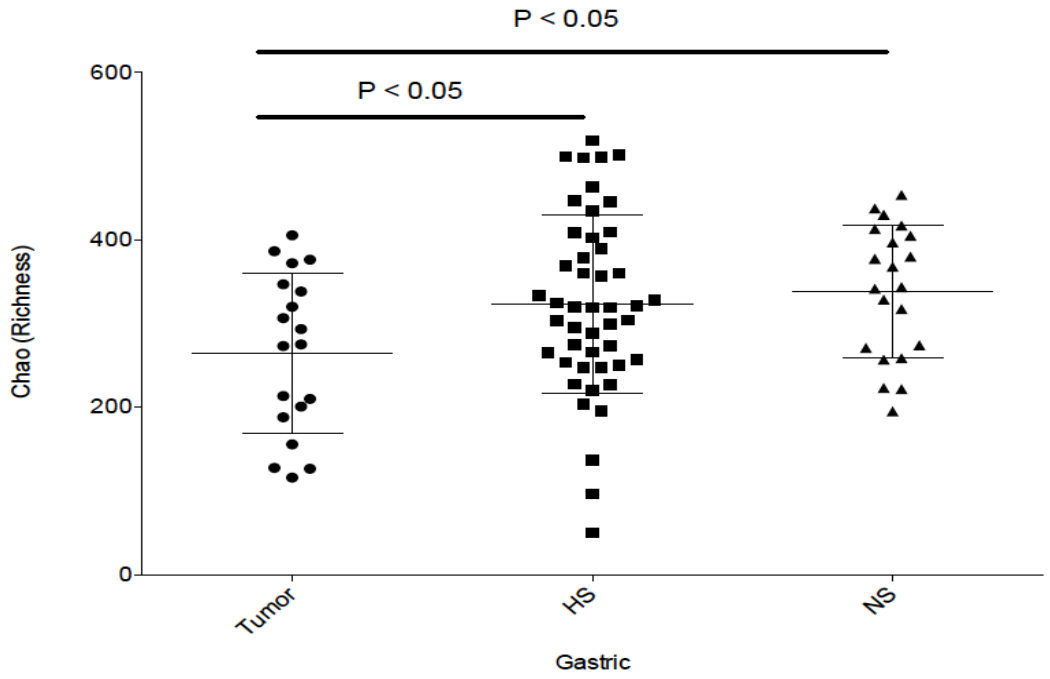

Figure 16:Bacterial richness level comparing gastric cancer samples (tumour) with normal stomach samples (NS) from the same cancer patient and healthy control samples (HS).

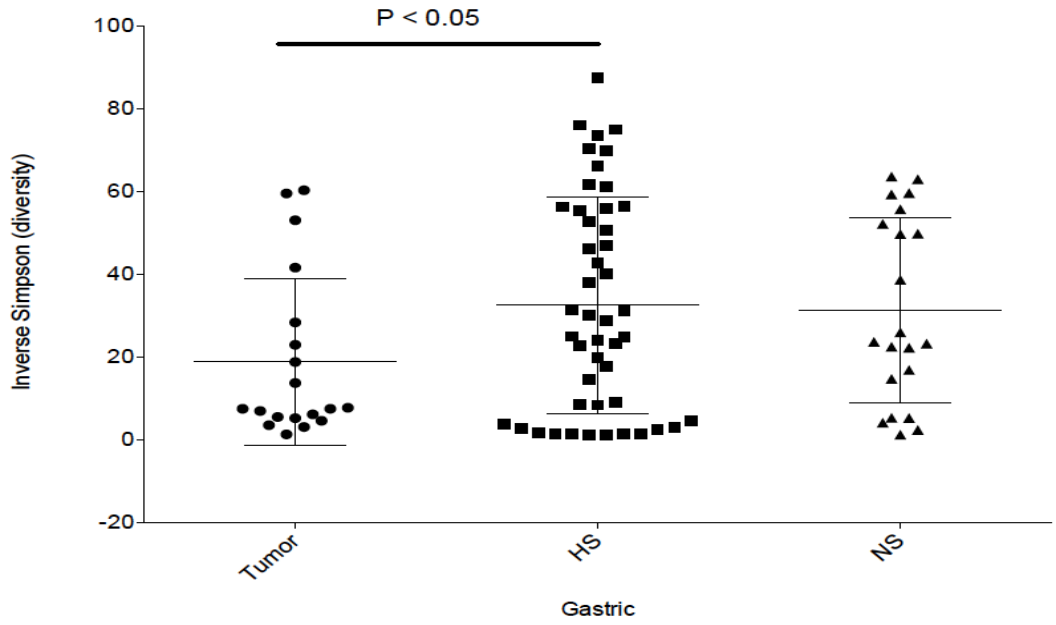

Figure 17:Bacterial diversity level comparing gastric cancer samples (tumour) with normal stomach samples (NS) from the same cancer patient and healthy control samples (HS).

Amongst the five bacterial phyla present in gastric cancer patients (Actinobacteria, Bacteroidetes, Firmicutes, Fusobacterium and Proteobacteria) there was an observed increase in the relative abundances of 12 bacteria at the family level of tumour samples (n=20) when compared with matched normal gastric mucosal samples of the same cancer patient. Streptococcaceae (p=0.002), Actinomycetaceae (p=0.001), Micrococcaceae (p=0.012), Eubacteriaceae (p=0.001), Staphylococcaceae (p=0.016), Fusobacteriaceae (p=0.007) and Leptotrichiaceae (p=0.012) were more abundant in tumour samples compared to adjacent normal stomach. There was however a decrease in the relative abundances of the family Helicobacteraceae in the same group of tumour samples (Figure 18).

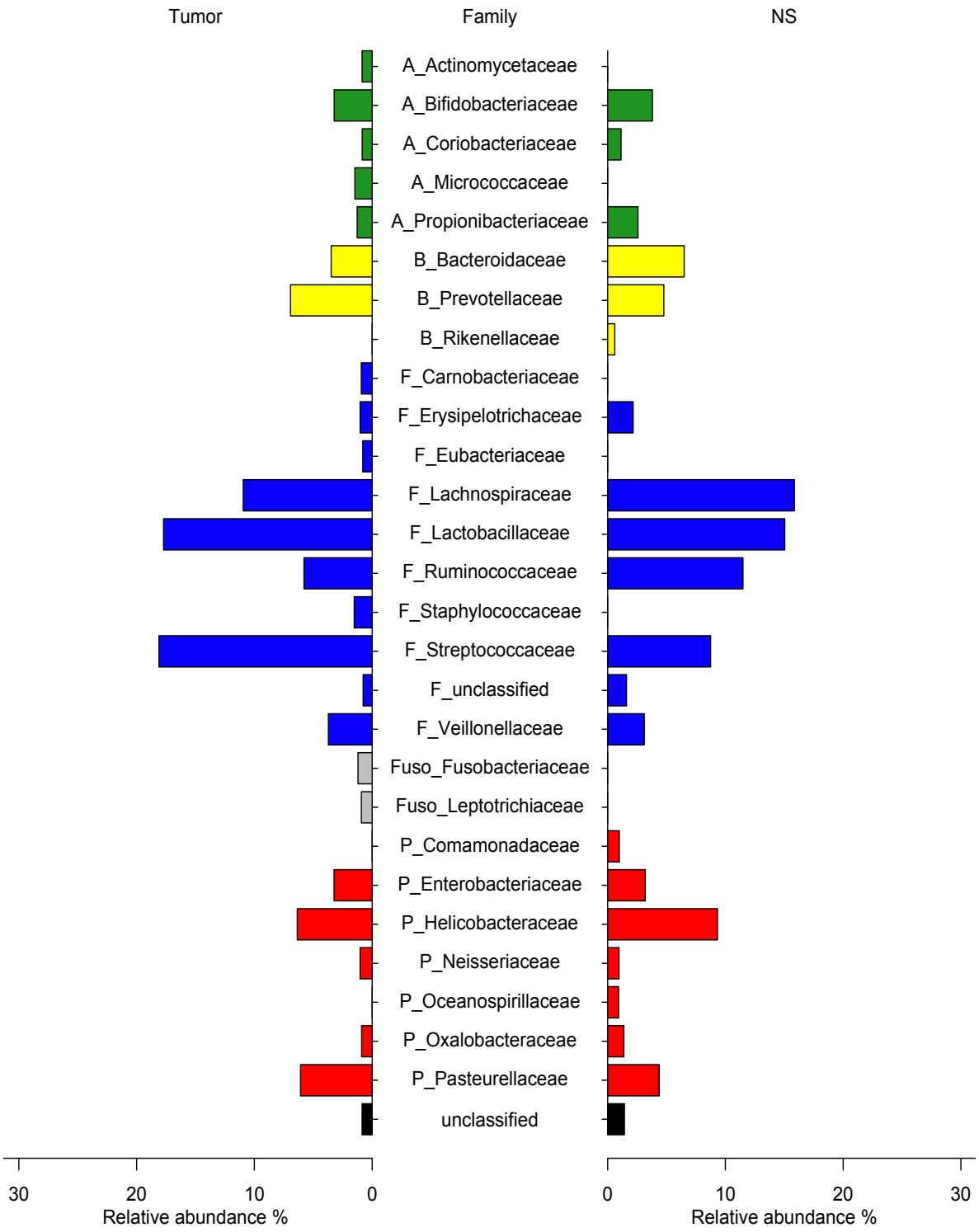

Figure 18:Bacterial family level assessment comparing gastric cancer (tumour) samples with normal stomach samples from the same cancer patient (NS) demonstrating the variations of bacteria between both groups.

When comparing the family level of tumour samples (n=20) versus healthy control patients (n=55), the results demonstrated significantly higher diversity of Lactobacillaceae (p=0.001), Streptococcaceae (p=0.009), Eubacteriaceae (p=0.001), Staphylococcaceae (p=0.001), Leptotrichiaceae (p=0.001), Enterobacteriaceae (p=0.04) and Pasteurellaceae (p=0.03) as shown in Figures 19,20 and 21. The predominant species in tumour samples from the family level, compromised gut microflora commensals (Lactobacillus fermentum, Escherichia coli) and from human oral and respiratory tracts (Streptococcus salivarius, Streptococcus mitis). Interestingly there were also a few other opportunistic pathogens that were upregulated in cancer samples, these include Streptococcus anginosus and Klebsiella pneumonia and Clostridium. The composition of *H pylori* containing Epsilonproteoobacteria class was most prevalent in control group (27%). However by analysing the *Helicobacter* group family, the gastric cancer group demonstrated a significant decrease of relative abundance of microorganisms from the Helicobacteraceae family. This observation is consistent with previous findings suggesting that *H Pylori* could potentially modify the gastric microbiota, however does not significantly affect its diversity[105]. Eradication of *Helicobacter Pylori* has been shown by previous studies to reduce the incidence of gastric cancer [97] however such eradication remains controversial with the survival of other genotoxic bacteria in an altered gastric environment with lower acidity [48].

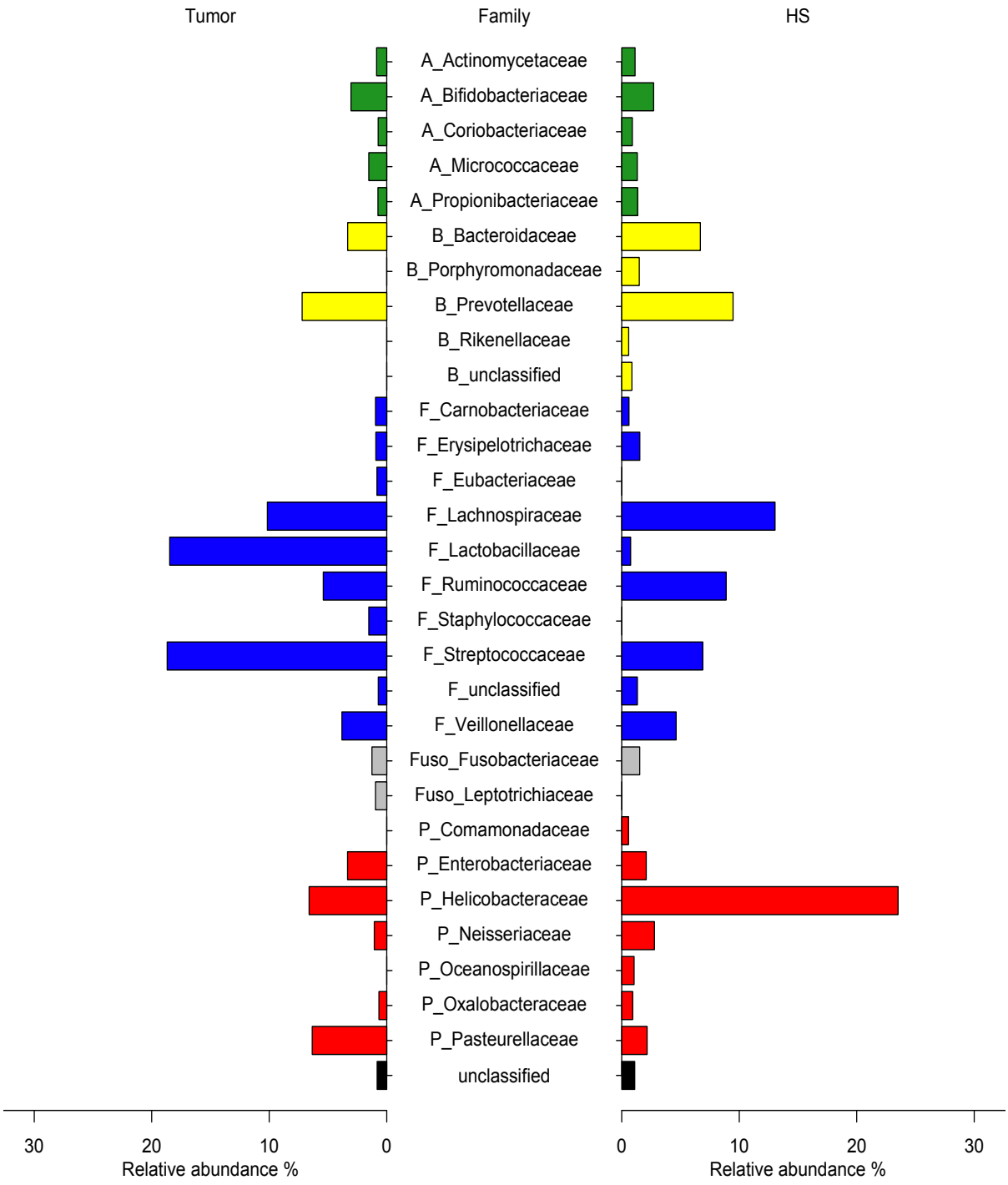

Figure 19:Bacterial family level assessment comparing gastric cancer samples (T) with samples from control patients (HS) demonstrating more prevalence of Lactobacillaceae and streptococcaceae in tumour samples. Relative abundance of *Helicobacter* is significantly higher in control patients.

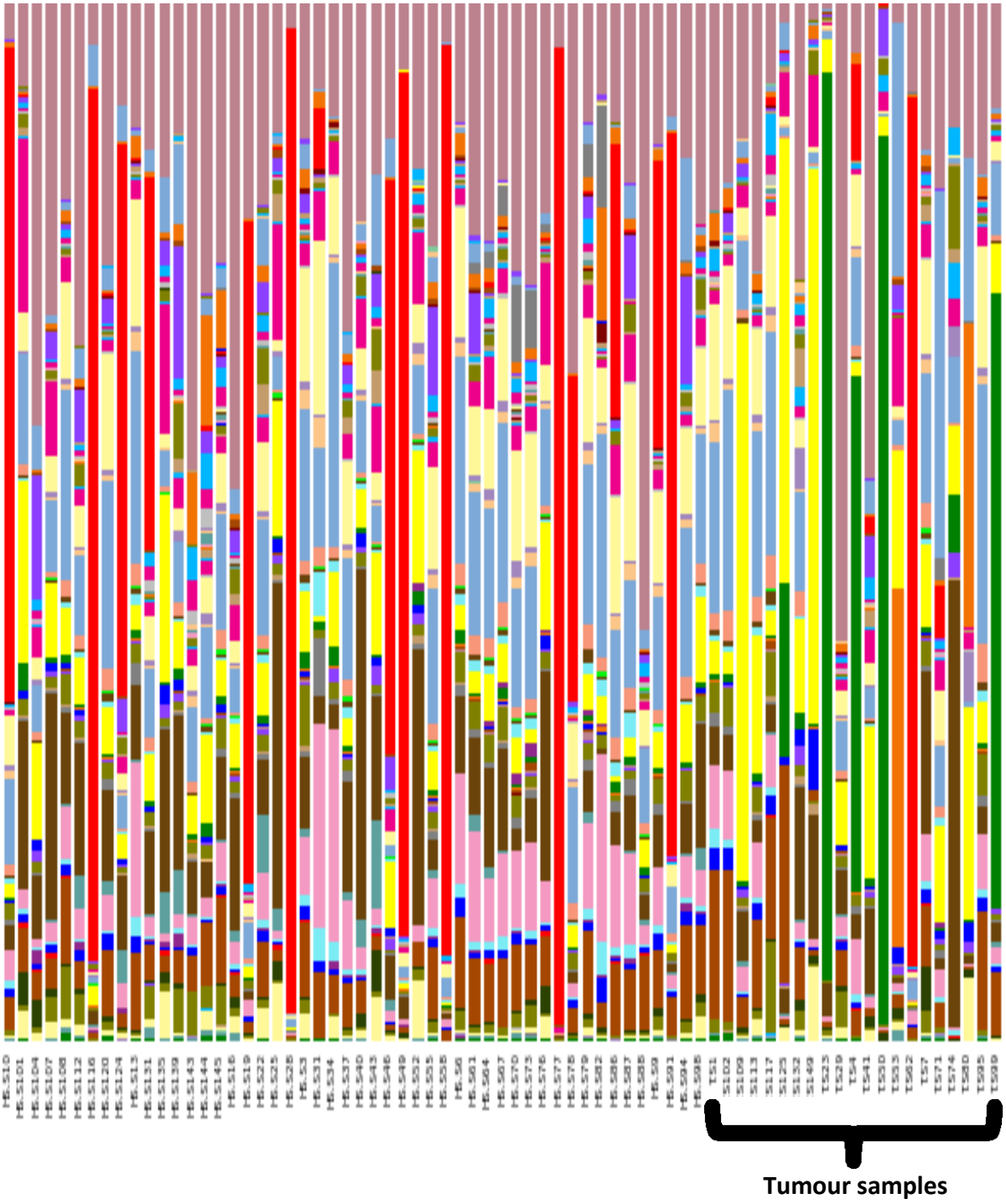

Figure 20:Bar chart demonstrating the level of variation between bacterial species in gastric cancer samples (tumour) with samples from control patients (HS) showing an increase of lactobacillaceae (green colour) and Streptoccaceae (dark yellow).

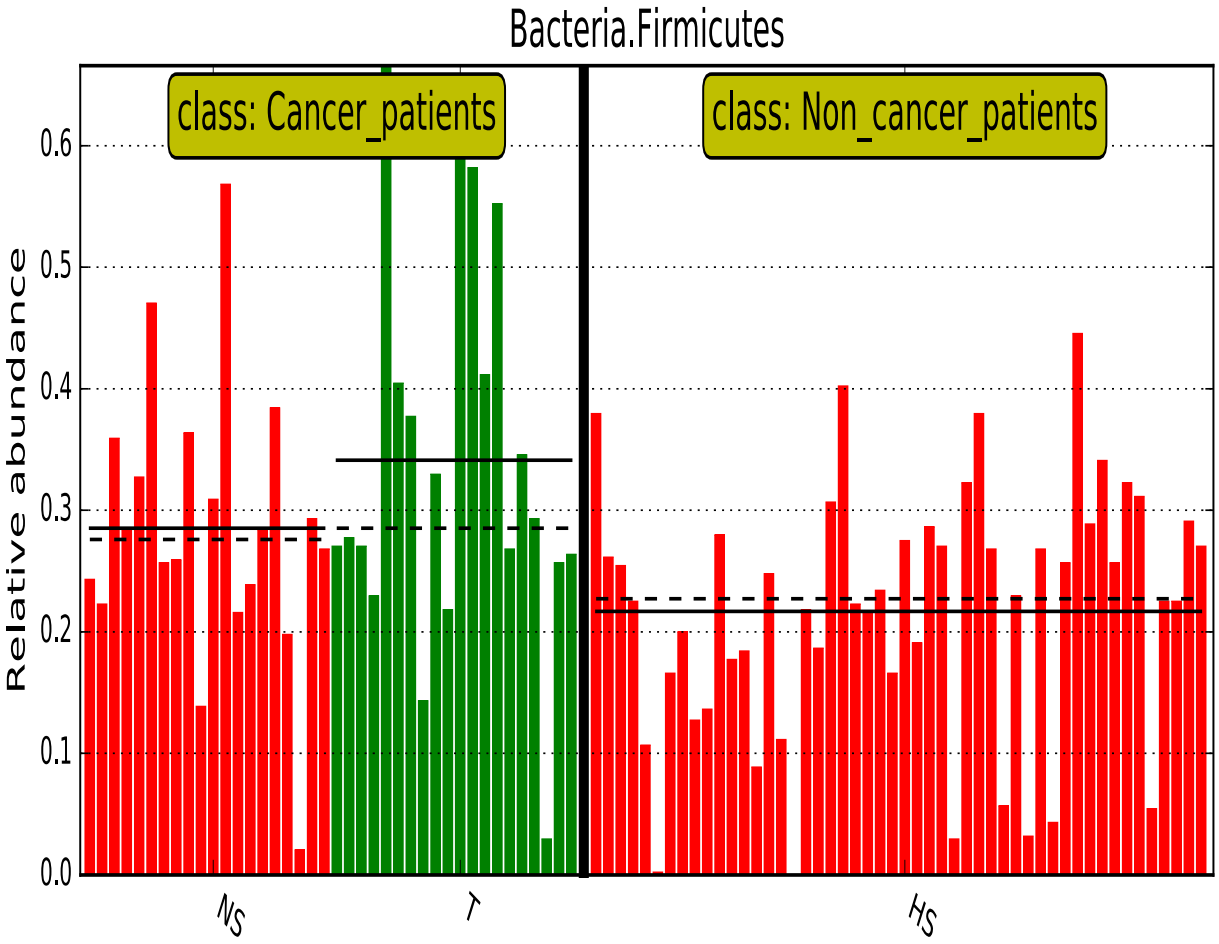

Figure 21:Bacterial phylum level assessment comparing gastric cancer samples (T) with normal samples from the same cancer patient (NS) and control patients (HS), demonstrating an overall increase of Firmicutes in T and NS when compared to HS.

6.3.3.1.2 Oesophageal cancer 16S rRNA sequencing analysis

The microbial communities of oesophageal mucosa from 60 patients, including 30 patients with oesophageal cancer and 30 healthy controls, were analysed using high-throughput sequencing 16s rRNA gene. The Chao 1-estimated richness and Inverse Simpson in oesophageal cancer patients (210.7,18.2) was significantly lower than normal stomach samples from the same cancer patient (295.6, $p < 0.05$, 38.6, $p < 0.01$ respectively). However there were no significant differences when comparing cancer with normal oesophageal samples from the same cancer patient (212,18.6) and healthy controls (224, 20.6). In addition, normal oesophageal samples from the same cancer patient showed significantly higher diversity and richness ($p < 0.05$, $p < 0.01$) compared to normal oesophageal samples from the same cancer patient, while only significant diversity when compared to control patients. Figure 22 and 23 below show the bacterial richness and diversity between groups.

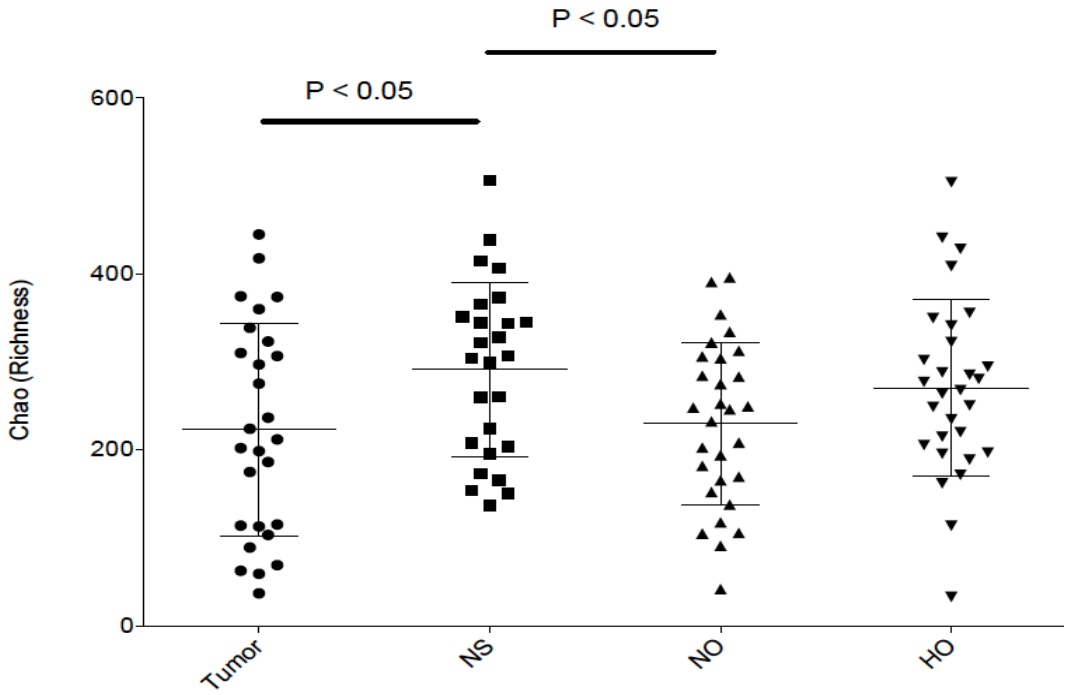

Figure 22:Bacterial richness level comparing oesophageal cancer samples (tumour) with normal stomach samples from the same cancer patient (NS), normal oesophageal samples from the same cancer patient (NO) and healthy control samples (HO).

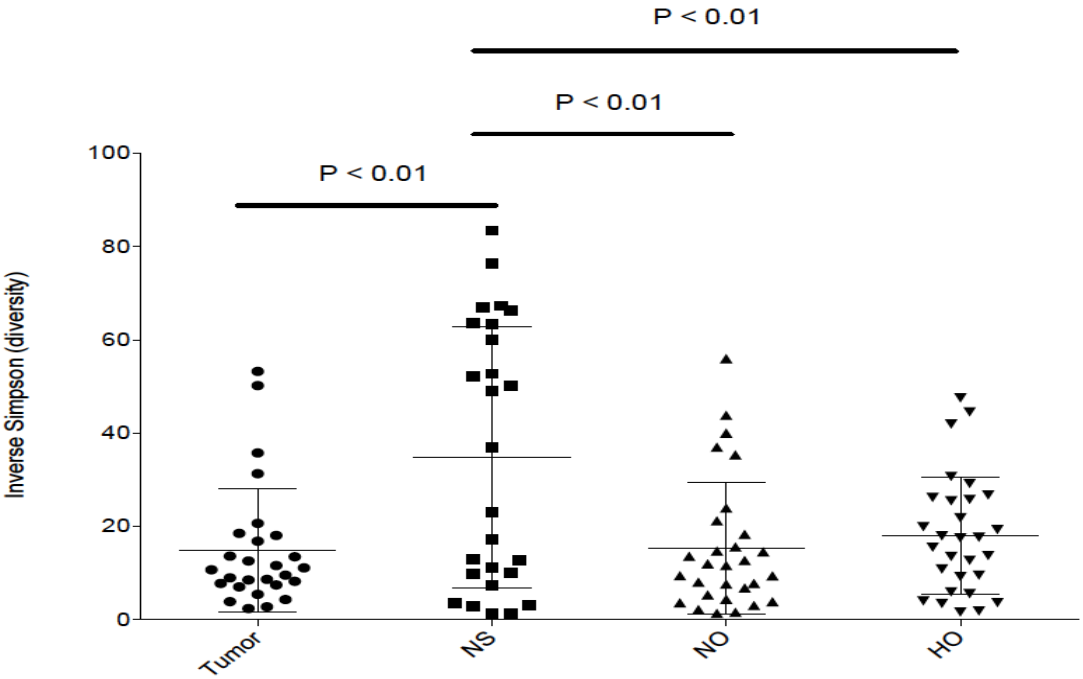

Figure 23:Bacterial diversity level comparing oesophageal cancer samples (tumour) with normal stomach samples from the same cancer patient (NS), normal oesophageal samples from the same cancer patient (NO) and healthy control samples (HO).

Amongst the 5 bacterial phyla present in oesophageal cancer specimens (Actinobacteria, Bacteroidetes, Firmicutes, Fusobacterium and Proteobacteria) there was an observed increase in the relative abundances of 15 bacteria at the family level of cancer samples (n=30) when compared with matched normal gastric mucosal samples of the same cancer patient. Following the same trend as the microbial composition of gastric tumours, Streptococci were also the most prevalent between the samples (p=0.003) and a decrease in the relative abundances of the family Helicobacteraceae in the same group.

However when comparing the family level of bacteria present in oesophageal tumour samples (n=30) versus healthy control patients (n=30) there was no changes in the relative abundances of bacteria, apart from micrococci which belongs to the Actinobacteria phyla, and was only found to be significantly higher in control samples (p =0.012). Family level assessment between groups is shown in Figures 24 and 25 below.

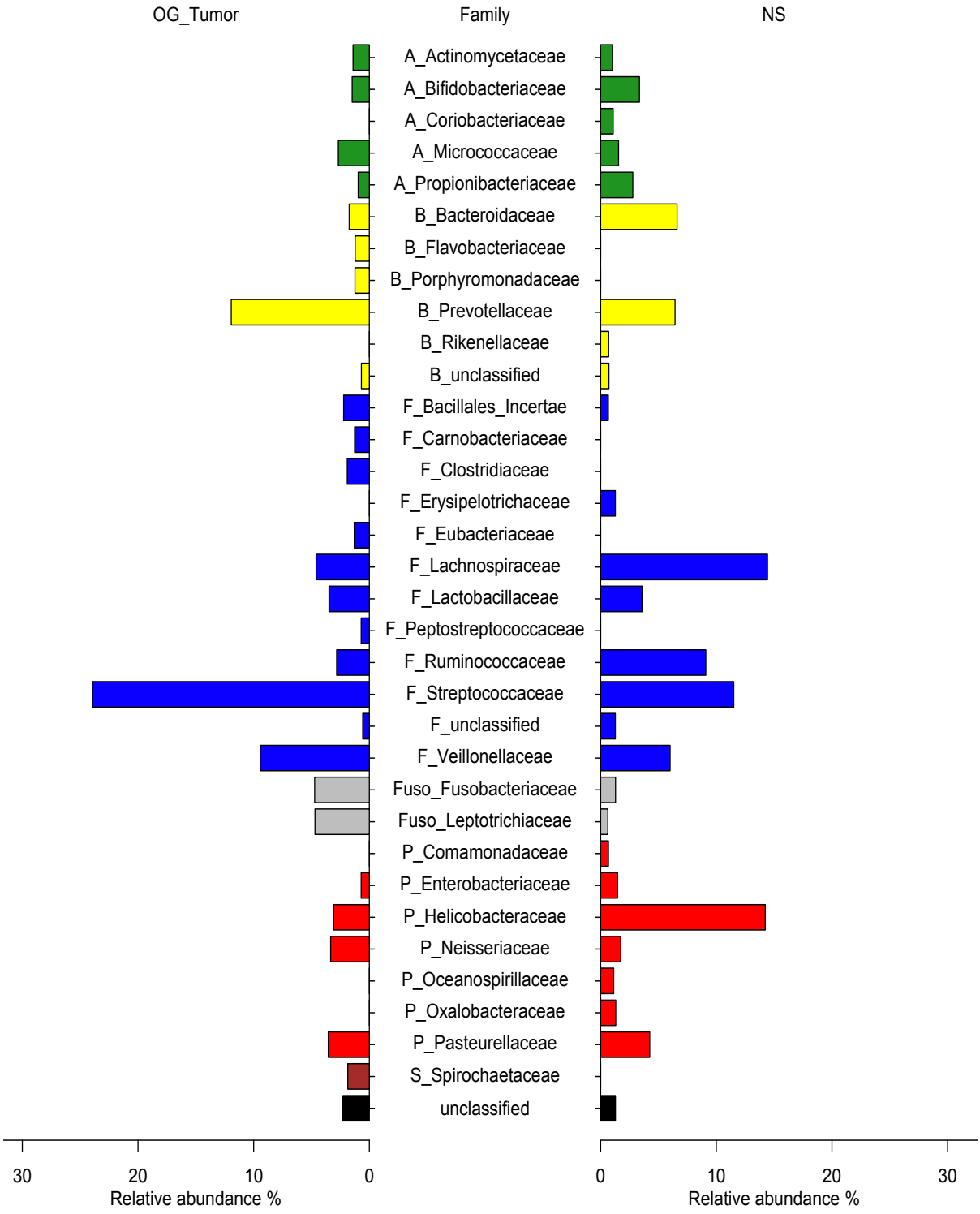

Figure 24: Bacterial family level assessment comparing oesophageal cancer samples (tumour) with normal stomach samples from the same cancer patient (NS), demonstrating a significant increase of mainly Streptococcus in tumour samples.

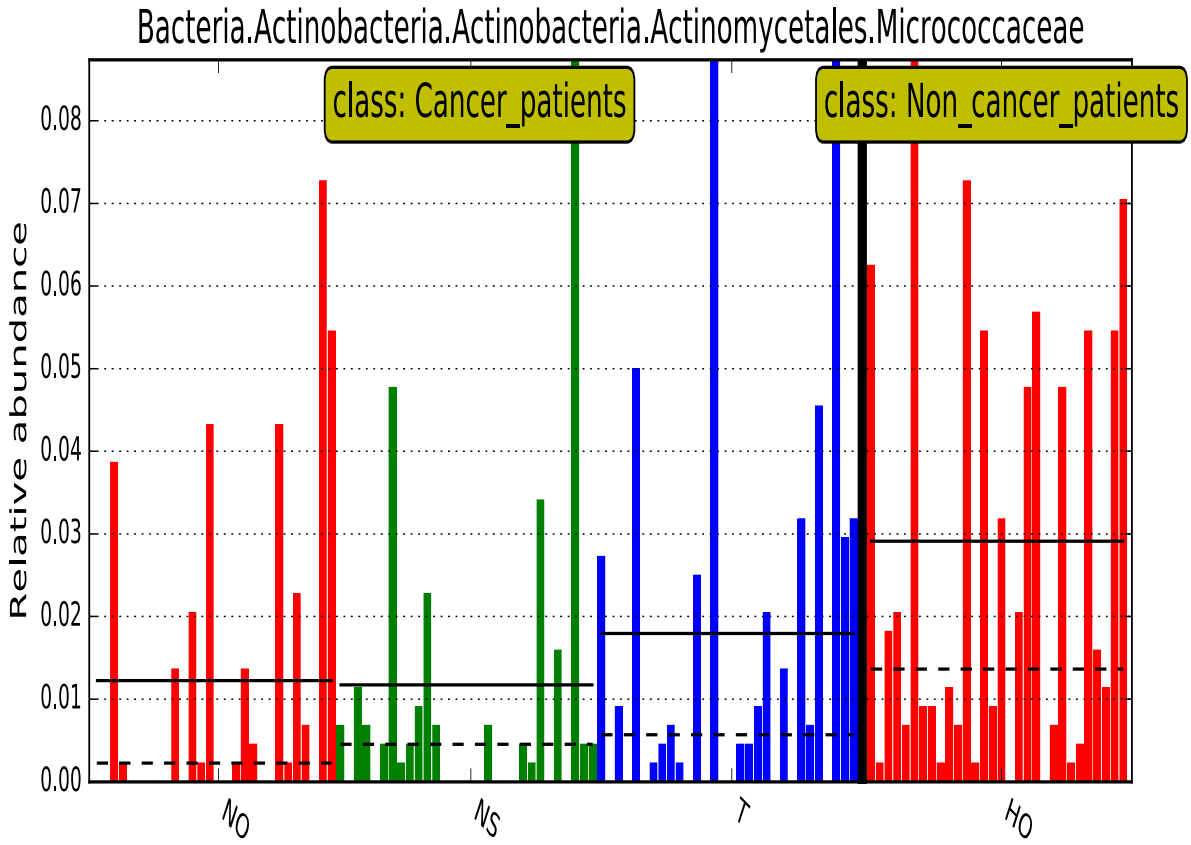

Figure 25: Bacterial family level assessment comparing oesophageal cancer samples (T) with normal stomach samples from the same cancer patient (NS), normal oesophageal samples from the same cancer patient (NO) and healthy control samples (HO), demonstrating a significant increase of Micrococcaceae in HO compared to the rest of the groups.

6.3.3.2 Results of post-chemotherapy analysis

There are numerous downstream consequences of antineoplastic agents on the gut microbiome, an effect that is suggested to influence drug responses. In this experiment, the microbial communities of gastric mucosa from 14 patients were compared before and after the administration of chemotherapy using high-throughput sequencing of the 16s rRNA gene. The data shown demonstrate variations in the bacterial diversity that occurs after administration of chemotherapy. From those bacteria, the family group that showed significant reduction in abundance were from the genera of Streptococci ($p<0.024$),

Prevotella (p<0.019) and Fusobacteria (p<0.032) after having chemotherapy treatment. Results are shown in Figure 26 and 27 below.

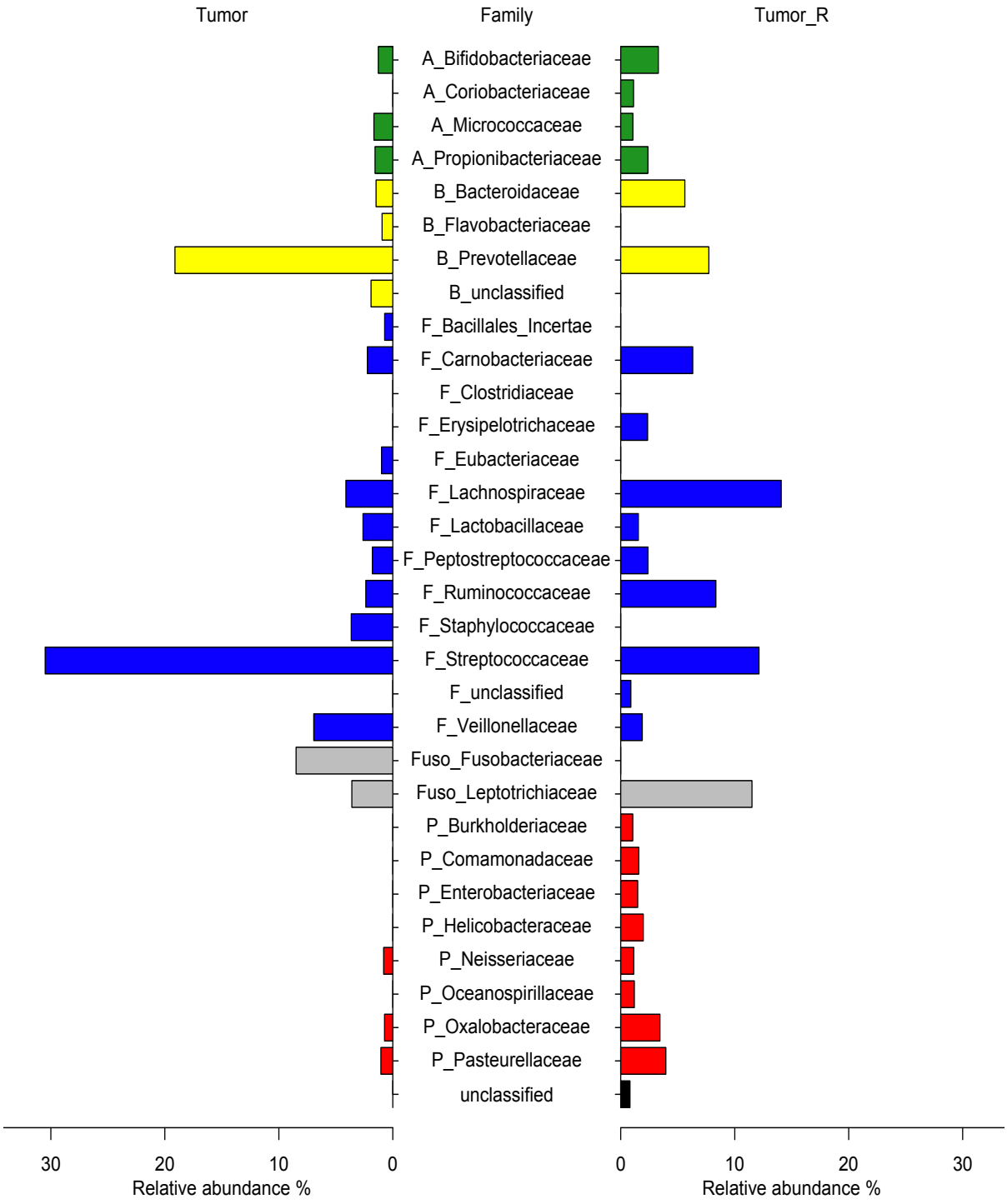

Figure 26:Comparison of bacteria between pre-chemotherapy patients (tumour) and post-chemotherapy (tumour_R) showing a significant decrease in Streptococci, Prevotella and Fusobacteria after having chemotherapy.

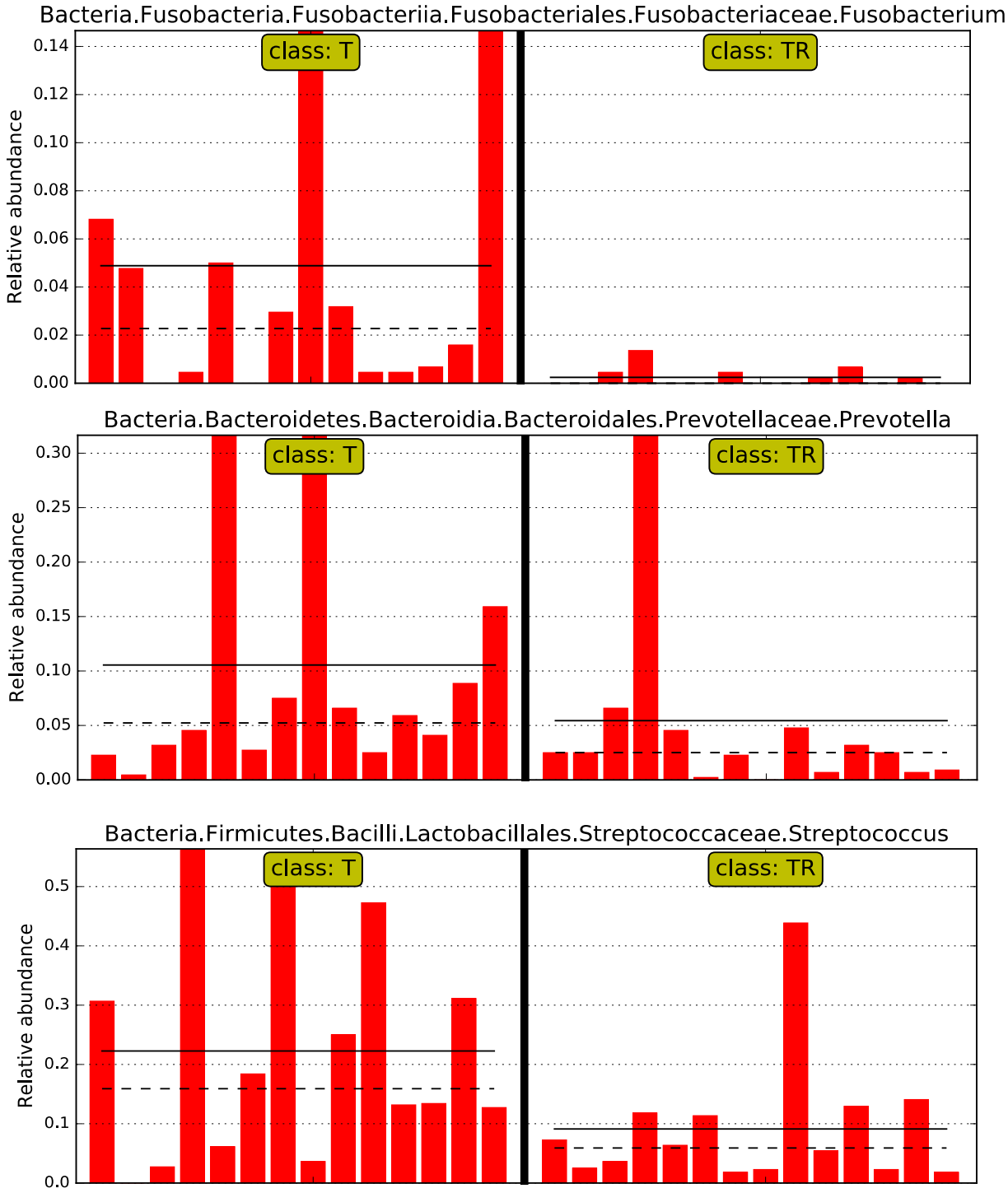

Figure 27:Comparison of bacteria between pre-chemotherapy patients (T) and post-chemotherapy (TR) showing a significant decrease in Streptococci, Prevotella and Fusobacteria after having chemotherapy. Dotted line represents the median value and full line is the mean.

6.3.4 Discussion

The inhospitable acidic environment of the stomach acts as a barrier killing many microbes in the gastrointestinal tract. However, due to the physiological changes in the stomach from several factors including acid-reducing therapy or during gastric cancer progression, this provides an opportunity for other foreign microbes to colonize the gastric mucosa This study has characterised the tumour associated microbiome in patients with oesophago-gastric cancer. Although there is a substantial variation in the microbiome between individuals, the principal finding of 16S rRNA gene sequencing analysis was the observation of a higher prevalence of Firmicutes (Lactobacillaceae Streptococcaceae, Staphylococcaceae, Eubacteriaceae families), Proteobacteria (Escherichia coli, Klebsiella, pneumonia, Pasteurellaceae), Fusobacteria (Fusobacteriaceae, Leptotrichiaceae) and lower abundance of *H pylori* in tumour samples.

Lactobacilli are gram-positive, rod-shaped and micro-aerophilic bacteria with some similarities to *H pylori*. The characteristic feature of Lactobacillus metabolism is the conversion of lactose to lactic acid, thus leads to acidification of the bacterial environment consequently resulting in acidification of the gastric mucous layer [106]. Lactobacilli are able to acclimatise to the acid environment and colonize the stomach due to their acidophilic properties[107]. Furthermore, some species of Lactobacilli release urease enzyme with an optimum activity at pH 3 to 4, similar to that of *H pylori*[108]. Acidification of the gastric antral mucosa can modulate gastric physiology by the inhibition of gastrin and a reduction in gastric acid secretion[109]. Thus, acid-generating Lactobacilli near to the surface of the antral gastric epithelium could potentially decrease gastric acid secretion. In comparison with *H pylori*, the bacteria alkalinises the gastric antral mucosa, which in turn increases gastrin and

consequently also acid secretion[110]. The impact of Lactobacilli colonizing the gastric mucosa may however be different in different regions of the stomach.

The data has also shown that bacteria more frequently colonize oesophagi-gastric tumours compared to the surrounding mucosal tissues, and the most frequent colonizers are gram-positive cocci such as Streptococci [111]. The most commonly isolated bacteria from stomach and oesophageal cancer were streptococci, which are known to have a role in pathogenesis of human disease, including cancer. From those species, Streptococcus Bovis has been reported in a previous study to have an association with the development of colorectal cancer by assessing the potential of the bacteria to produce pre-inflammatory properties in a rat model leading to colorectal carcinogenesis[112].

Regarding the homogeneity of the microbiota in different areas of the stomach and oesophagus, some have found no variations in microbiota of the antrum and corpus of their populations, however others reported a difference[93], which remains unclear. More studies involving the microbiota-host-environment interactions, including the effect of diet, gender, geographic region and acid suppression medications are needed to fully understand the role of oesophago-gastric bacteria in human health and disease.

The overall results correlate with a previous study by Dicksved *et al*, characterizing the differences in gastric microbial composition among patients with gastric cancer and healthy controls using a different method for molecular profiling (terminal restriction fragment length polymorphism) [48]. Although the numbers of patients were few (10 cancer and 5 controls) they have shown a significant rise in Firmicutes phyla similar to what is presented

from this study using more patient numbers. Several studies have reported changes in a variety of diseases in the oesophagus including Gastro-Oesophageal Reflux Disease (GORD) and Barrett's Oesophagus. A study by Elliott *et al.* demonstrated a similar trend with reduction of microbial diversity in oesophageal adenocarcinoma when compared with samples with Barrett's Oesophagus and control patients using a Cytosponge™ device as their method of collecting samples[53].

Chemotherapy has got the potential to damage the mucus layer and disrupt the intestinal barrier allowing some microbiota to penetrate the mucosa leading to systemic infections and microbial alteration. Fusobacteria has shown a marked decrease in its abundance after giving chemotherapy to gastric cancer patients[113]. In addition to its presence in oesophago-gastric tumour samples, they showed lower abundance to adjacent normal muocsa from the same cancer patient. Fusobacteria, anaerobic gram-negative rods, present in oesophago-gastric mucosal tissue have recently been reported to be associated with the development of colorectal cancer[113]. In their study Jing *et al.* explored the effect of chemotherapy at different stages using faecal samples of colorectal cancer patients to study the complex interactive network on intestinal microbial community. During different stages of the treatment, colorectal cancer patients had unique ecological network models than non-chemotherapy samples. Firmicutes, bacteroidetes and fusobacteria dominated during each stage of chemotherapy treatment [114]. Further studies have also been able to demonstrate the ability of fusobacteria to colonise colonic mucosa and stimulate oncogenic gene expression[115].

Some investigators have suggested that the gastric microbiota is involved in the production

of carcinogens through the promotion of inflammation[116]. However, it is not clear if bacterial overgrowth is a consequence of the stepwise progression to cancer by producing an environment that favours bacterial proliferation. In a recent study including 212 patients with chronic gastritis and 103 with gastric cancer, the amounts of bacteria between both groups was measured using quantitative PCR demonstrating an increase in bacterial load in patients with gastric cancer as compared with those with chronic gastritis [117]. Nevertheless, further research is required to clarify the mechanisms by which these changes occur, and the possible relationship to cause or effect. Other data suggest that differences in the gastric microbiota might be responsible for a higher prevalence of gastric cancer in some geographic regions.

6.3.5 Conclusion

This experiment has established the upregulation of specific bacteria with the potential to develop oesophago-gastric cancer by inducing the carcinogenic pathways, the stimulation of cell proliferation and production of several metabolites. These metabolites may include bacterial enzymes, which can provide a potential mechanism of VOC production within oesophago-gastric cancer patients.

This study describes the characteristic microbiota of oesophago-gastric cancer providing potentially important insight in to tumorigenesis and challenging the long-held concept of sterility within the stomach.

6.4 Quantification of tyrosine phenol lyase enzyme in fusobacteria

6.4.1 Background

Results presented in the previous section demonstrate the predominant bacteria in oesophago-gastric cancer and affords an opportunity to explore their role in the release of specific cancer associated VOCs. There is evidence that intestinal bacterial may play a role in phenol production in oesophago-gastric patients via bacterial action upon tyrosine found within the gastric juice[118]. The role of the colonic microbiota in phenol production from tyrosine has long been established[82]. Phenol production is linked to aromatic amino acid metabolism through the process of proteolytic fermentation. Specifically, phenylalanine hydroxylase catalyses the conversion phenylalanine to tyrosine and tyrosine phenol-lyase is involved in the conversion of tyrosine to phenol.

There are several studies investigating phenol concentrations in the biofluids of cancer patients. Kumar *et al.* showed that cancer patients had a higher median concentration of methyl phenol compared to patients with other upper gastrointestinal disease and healthy controls[23]. In addition, Phenol has been also demonstrated to be present at high concentrations ($p<0.001$) within exhaled breath of oesophago-gastric cancer patients[21].

Species level analysis from the previously described study (above) revealed the presence of a number of bacteria, which are known to contain the tyrosine phenol-lyase enzyme (Fusobacteria, Clostridium, and Escherichia). The presence of these bacteria within the stomach microbial community indicates that there is the potential for tyrosine to be converted to phenol within the stomach via bacterial action. The identification of bacteria

with the potential to facilitate this reaction has provided the initial evidence to establish that stomach bacteria have the potential to modulate the conversion of tyrosine to phenol within the stomach.

6.4.2 Aims

This experiment aimed to utilise qPCR to identify the tyrosine *phenol lyase* (TPL) enzyme gene present in Fusobacterium species, from DNA extracted from human oesophago-gastric cancer tissue samples.

6.4.3 Methods

Replicates of extracted DNA samples from the same gastric tissue performed in the previous section were provided (Section 1.3.2.1.2). Concentrations were measured using the Qubit DNA High Sensitivity Assay Kit (Thermo Fisher Scientific, Cat No: Q32854), following the recommended guidelines provided by the manufacturer.

6.4.3.1 *PCR primers*

PCR primers were designed to identify a range of regions across the TPL gene, in six groups of degenerated primers, listed in Table 11. Details of the design were listed in Figure 28. Primers were synthesised by Sigma Aldrich.

Forward primer:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A			8		52			52	34			9		52				14		
T			22			28						13	52		33			10		
G	52	52		52			52		18	49						52			52	52
C			19			24					52	26			19		49	22		
Fusobact. sequence	G	G	T	G	A	T	G	A	A	G	C	Y	T	A	T	G	C	W	G	G
Consensus	G	G	H (Y)	G	A	Y	G	A	R	G	C	H (Y)	T	A	Y	G	C	H	G	G

Reverse primer:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A			6			13	29					7			11	11		10		52
T			13			18		52	5			19			17			12		
G	52	52		52	52	8			43	48		5	52	52	11		52	11	52	
C			30			13	19				52	21			13	41		19		
Fusobact. sequence	G	G	W	G	G	A	C	T	T	G	C	T	G	G	W	A	G	A	G	A
Consensus	G	G	H	G	G	N (H)	M	T	K	G	C	N (M)	G	G	N	M	G	N	G	A

Figure 28:Illustrating the degenerated primers from designed primer sequences.

Primer Group	Primer Number	Primer Name	Sequence	Tm	Expected PCR Product Size (bp)
A	8471	Tpl_loc1_For	GTTCCTACTCACCAAGGAAG	57.8	800
	8472	Tpl_loc1_Rev	GCTCAAGCTCTTGCAGCAGC	67.8	
B	8473	Tpl_loc1_deg1_For	GTNCCNACYCAYCARGGVMG	46.8	800
	8474	Tpl_loc1_deg1_Rev	KCDGCNGCNARNGYYTGNGC	60.6	
C	8475	Tpl_loc1_deg2_For	GTNCCBACYCACCARGGVAG	50.0	800
	8476	Tpl_loc1_deg2_Rev	KCKGCNGCHAGRGYYTGNGC	64.7	
D	8477	Tpl_loc2_deg1_For	GGTGATGAAGCYTATGCWGG	71.1	700
	8478	Tpl_loc2_deg1_Rev	TCTCTWCCAGCAAGTCCWCC	57.3	
E	8479	Tpl_loc2_deg2_For	GGHGAYGARGCHTAYGCHGG	72.9	700
	8480	Tpl_loc2_deg2_Rev	TCNCKNCCNGCMAKNCCDCC	62.4	
F	8481	Tpl_loc2_deg3_For	GGYGAYGARGCYTAYGCHGG	72.9	700
	8482	Tpl_loc2_deg3_Rev	TCNCKNCCKGCMAKDCCDCC	62.4	

Table 11**:**Degenerated Primers for TPL Amplification, TM (primer melting temperature).

6.4.3.2 *Initial PCR under standard conditions*

PCR was performed using standard protocols (OneTaq mastermix with standard buffer, New England Biolabs, Cat No: M0482S). Reaction conditions are detailed in Tables 12 and 13. A standard volume of 0.5 µL DNA plus 49.5 µL PCR mastermix solution was used for each sample. Primer groups were segregated according to the lowest annealing temperature within the pair. Primer groups A, D, E and F were annealed at 50°C. Primer groups B and C were annealed at 40 °C.

Reagent	**Volume (µL)**
Onetaq mastermix (2x)	1000
Forward Primer (10 µM)	40
Reverse Primer (10 µM)	40
Water	900

Table 12:OneTaq PCR Mastermix Solution.

Temperature (°C)	**Time (seconds)**	**Number of Cycles**
94	30	1
94	15	40
Annealing	60	
68	60	
68	300	1
4	Hold	-

Table 13:OneTaq PCR Cycling Conditions.

PCR reactions were analysed by Tapestation 2200 (Agilent), using the P5000 tape, as described in the instruction manual. A representative gel (samples 1 and 2) is shown in Figure 29.

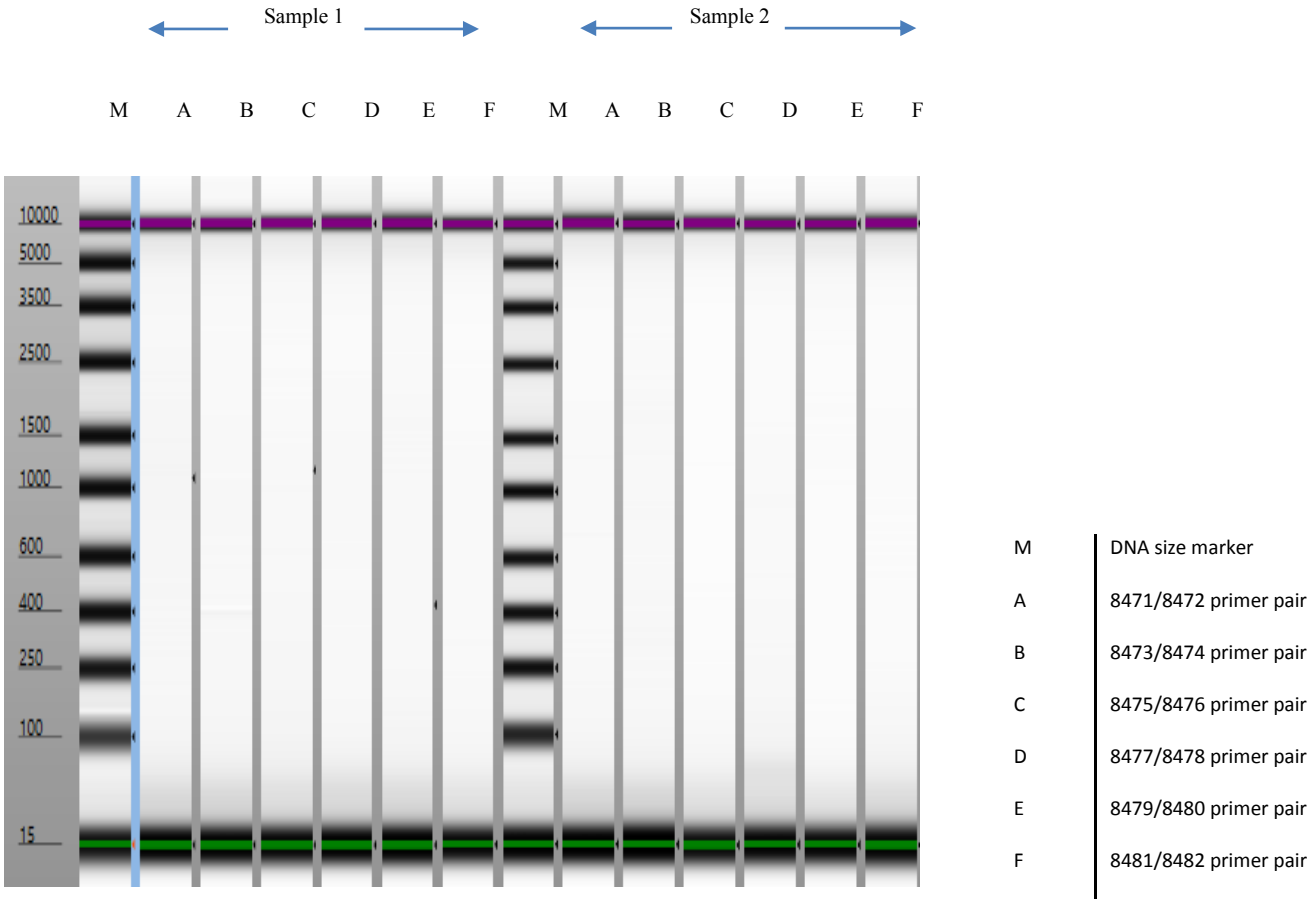

Figure 29:Tapestation analysis of samples.

No detectible bands were visible in any samples, for any combination of primer pairs. Likely causes for this were a DNA concentration that is too low. Additional DNA samples (and concentrations) were needed for further optimization of PCR conditions.

6.4.3.3 *Repeat PCR under optimised conditions and increased DNA concentrations*

To optimise the method and repeat the process, additional DNA samples with different concentrations were provided detailed in Table 14.

Sample Number (Tumour)	Sample Weight (g)	DNA conc. (ng/µl)	Total DNA conc.	Sample Number (Control)	Sample Weight (g)	DNA conc. (ng/ µl)	Total DNA conc.
1	0.001	7.1	35.5	1	0.004	13.4	67
2	0.008	9	45	2	0.005	29.4	147
3	0.007	33.2	166	3	0.004	15.3	76.5
4	0.003	17.7	88.5	4	0.003	11	55
5	0.005	31.5	157.5	5	0.007	7.4	37
6	0.008	16.4	82	6	0.006	59	295
7	0.004	7.9	39.5	7	0.007	18.3	91.5
8	0.003	14.8	74	8	0.020	9	45
9	0.006	63.1	315.5	9	0.006	8	40
10	0.004	11.4	57	10	0.012	42.8	214
11	0.001	45.3	226.5	11	0.007	15.5	77.5
12	0.010	62.7	313.5	12	0.005	10.6	53
13	0.002	7.5	37.5	13	0.006	27.2	136
14	0.018	184.2	921	14	0.006	17.9	89.5
15	0.004	8.2	41	15	0.007	9.3	46.5
16	0.004	27.7	138.5	16	0.006	38.4	192
17	0.008	23.4	117	17	0.006	17.7	88.5
18	0.003	12.6	63	18	0.006	36.1	180.5
19	0.002	28.8	144	19	0.008	14.8	74
20	0.001	19.2	96	20	0.003	19.5	97.5

Table 14:Additional DNA samples used to optimise the method.

Samples with the highest DNA concentrations were tested by PCR using an alternative polymerase (HotStarTaq Mastermix kit, Qiagen, Cat No: 203443), following the standard conditions recommended by the manufacture. Reagent volumes used are detailed in Table 15. Cycling conditions are detailed in Table 16. A standard volume of 2 µL DNA plus 48 µL PCR mastermix solution was used for each sample. Primer groups were paired in the same way as in the previous PCR steps.

Reagent	**Volume (µL)**
HotstarTaq Mastermix (2x)	175
Forward Primer (2 µM)	35
Reverse Primer (2 µM)	35
Water	91

Table 15:HotstarTaq PCR Mastermix Solution.

Temperature (°C)	**Time (seconds)**	**Number of Cycles**
95	900	1
94	60	35
Annealing	60	
72	60	
72	600	1
4	Hold	-

Table 16:HotstarTaq PCR Cycling Conditions.

PCR reactions were analysed by Tapestation 2200 (Agilent), using the P5000 tape, as described in the instruction manual. A representative gel for samples 9 and 14 was performed as shown in Figure 30 below.

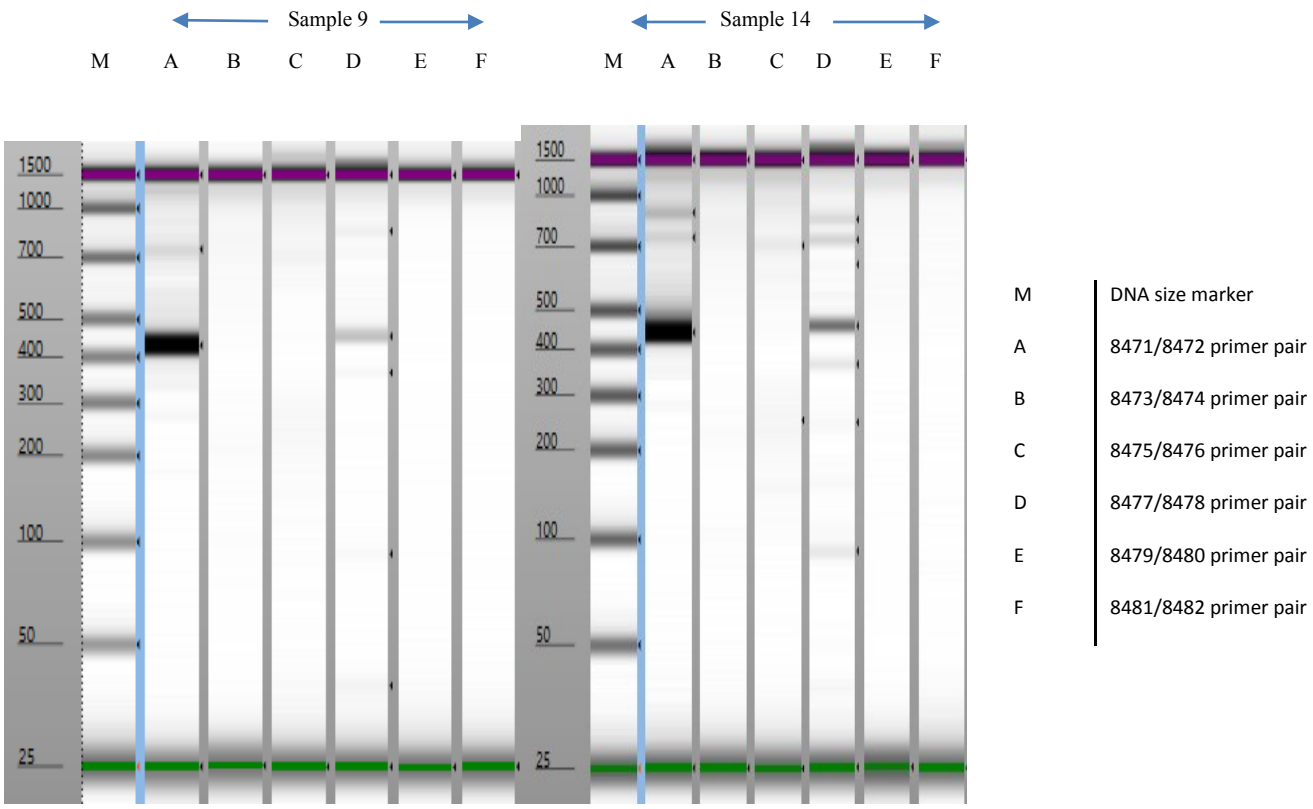

Figure 30:Tapestation analysis of samples 9 and 14 (HotStarTaq).

6.4.4 Results

In all samples tested in Table 14, a range of non-specific bands were observed in both primer group A and D. However, the dominant band was not the expected size (800 bp for primer group A, 700 bp for primer group D). Potential correctly sized bands may have been visible at very low concentrations. No bands were visible for the other primer groups (B, C, E and F) in any samples.

No decisive conclusion could be determined from this alone. Therefore to make sense of the banding pattern, gel extraction and sequencing analysis was performed on samples 9 and 14, primer groups A and D.

The remaining PCR reaction volume from these samples was separated on a 1 % agarose gel (running conditions, 120 volts, 45 minutes). The bands at 400 bp (the predominant band in both primer group A and D) were excised and extracted using the Monarch DNA gel extraction kit (New England Biolabs, Cat No: T1020S), following the manufactures guidelines. All samples were eluted with 10 µL elution buffer.

It was not possible to excise the 800 bp band from Primer group A, as the resolution is lower on agarose gels compared to the tapestation, meaning the band could not be suitably visualised. Concentration of the extracted bands is detailed in Table 17, as measured with the Qubit HS Assay Kit.

Sample	Concentration (ng/µL)
9, Primer Group A (400 bp)	155.0
9, Primer Group D (400 bp)	28.9
9, Primer Group D (800 bp)	11.9
14, Primer Group A (400 bp)	48.8
14, Primer Group D (400 bp)	48.4
14, Primer Group D (800 bp)	27.0

Table 17:DNA concentrations of the extracted bands.

These samples were analysed by sequencing (Source Bioscience) using the corresponding Forward PCR primer. Sequences were evaluated by BLASTn, results are summarised in Table 18. No hits were obtained relating to *Fusobacterium* in any of the samples. The BLASTn results are likely due to human DNA extracted from the tissue samples.

Sample	Primer	Blast Results
9 A 400 bp	8471	Human DNA sequence from clone RP1-18C9 on chromosome 20, complete sequence (97 % identification)
		Predicted Macaca fascicularis high mobility group protein B3 pseudogene (LOC107126591) (92 % identification)
9 D 400 bp	8477	Homo sapiens BAC clone CH17-155D14 from chromosome 7 complete sequence (99 % identification)
		Homo spaiens T-cell receptor gamma locus (TRG) on chromosome 7 (99 % identification)
9 D 800 bp	8477	No significant similarity
14 A 400 bp	8471	Human DNA sequence from clone RP1-18C9 on chromosome 20, complete sequence (90 % identification)
		Predicted Macaca fascicularis high mobiliy group protein B3 pseudogene (LOC107126591), misc RNA (89 % identification)
14 D 400 bp	8477	Homo sapiens BAC clone CH17-155D14 from chromosome 7 complete sequence (99 % identification)
		Homo sapiens T-cell receptor gamma locus (TRG) on chromosome 7 (99 % identification)
14 D 800 bp	8477	>100 hits, including Human, chimpanzee and Rhesus macaque, Sumatran orangutan

Table 18:Sequencing *Analysis Summary (BLASTn).*

6.4.4.1 *PCR optimisation*

Since the bands obtained and sent for sequencing were quite non-specific, some PCR optimisation was performed**.** The aim of the optimisation was to reduce the number of non-specific bands and increase amplification of the band of interest (700/800 bp). Sample volume was very limited, so the annealing temperature was varied.

HotstarTaq Mastermix solution was prepared in the same way explained in the methods section.

A gradient program of varying annealing temperature, ranging from 40°C to 70°C, increasing in 10°C increments, was performed on samples 62, 80 and 98 (those with the highest DNA concentration), detailed in Table 19.

Temperature (°C)	**Time (seconds)**	**Number of Cycles**
95	900	1
94	60	10
40	60	
72	60	
94	60	10
50	60	
72	60	
94	60	10
60	60	
72	60	
94	60	10
70	60	
72	60	
72	600	1
4	hold	-

Table 19:HotstarTaq PCR Cycling Conditions, Annealing Temperature Optimisation.

There was neither improvement in yield of the band of interest, nor a reduction in the non-specific bands shown in Figure 31. Insufficient human samples remained to either repeat or optimise the methodology. As there was no improvement in yield, sequencing was not performed on the optimised PCR samples.

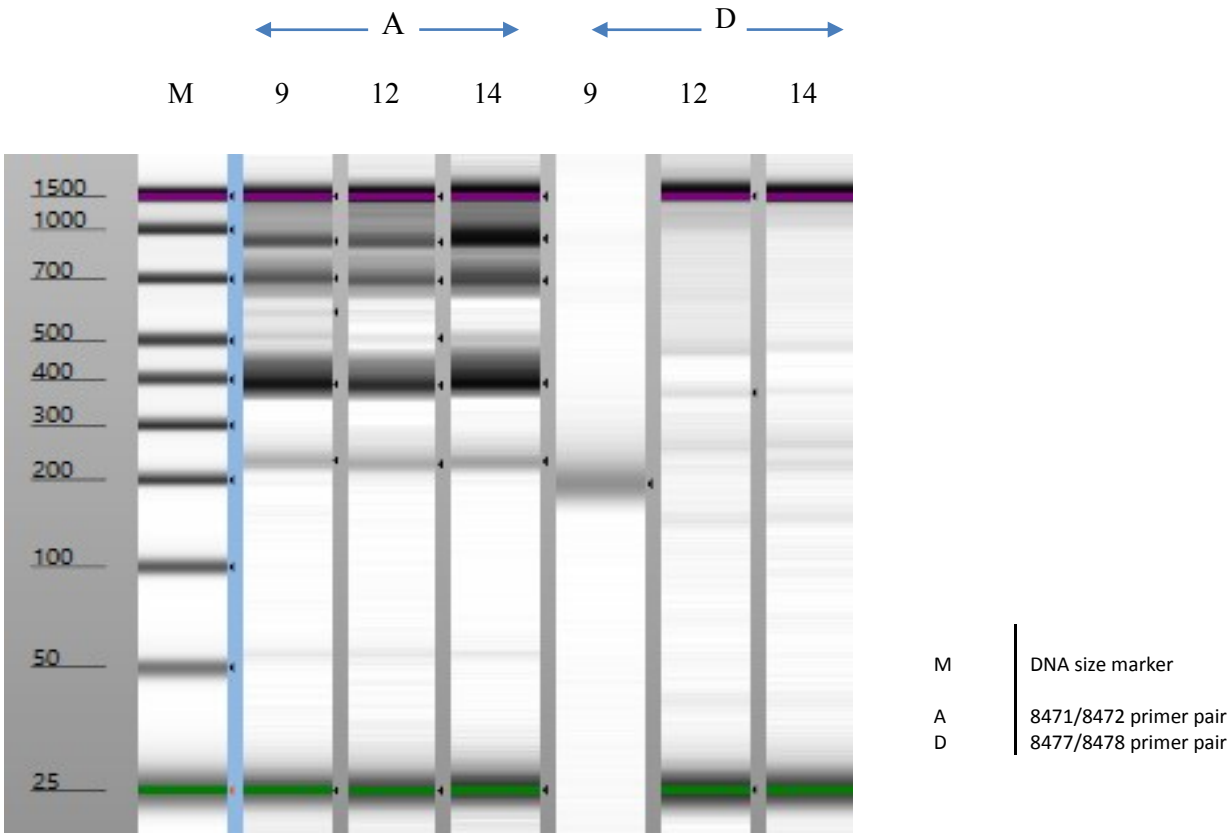

Figure 31:Tapestation analysis of samples 9, 12 and 14 (optimised PCR).

6.4.5 Discussion

Experiments yielded no clear evidence to confirm the presence of tyrosine *phenol lyase* gene either from the PCR or sequencing. The most likely cause of this was the unexpectedly low bacterial DNA yields from the tissue samples. The overall percentage of bacterial DNA in the samples was very small, due to limitation in the amount of tissue that could be obtained from patients. Performing PCR, particularly using the degenerated primers, with a very low concentration of target DNA has vastly reduced the odds of correct product formation, due to lack of available correct template DNA (bacterial DNA) in the total DNA concentration.

In order to increase the yield of bacterial DNA, culturing would be required in order to increase bacterial numbers prior to extraction of DNA. This could be achieved by removal of bacteria from the tissue, followed by growth on media specific for organism selection (e.g. Fastidious Anaerobe Agar – FAA, with horse blood), followed by harvesting of bacterial colonies and extraction of DNA. The PCR could then be repeated using enriched concentrations of bacterial DNA. Further PCR optimisation may also be required to increase specificity of the 700/800 bp products once the bacterial DNA is obtained.

6.4.6 Conclusion

Overall, based on these results alone, it is only possible to state that phenol producing bacteria appear to be present within the stomach. Whilst the target bacteria may contain the relevant gene, and have the potential to facilitate this reaction, further studies would also be required to definitively establish that they are biologically active.

7 BACTERIAL CULTURING

7.1 Background

From the work reported in the previous chapter, there is an emerging appreciation that oesophago-gastric cancers are associated with higher levels of certain bacterial species that may promote tumorigenesis. An increased bacterial presence within the stomach may provide a potential mechanism of elevated VOC production in patients with oesophago-gastric cancer, by bacterial action upon enzymes found within the gastric juice of these patients.

Recent studies have acknowledged that native human cells, including cancers produce VOCs [119].Symbiotic and pathogenic microbes may also contribute to VOC production within humans [120]. It is well established that microbes, such as bacteria and fungi, produce characteristic odours. These odours are a result of several metabolic processes leading to the production of characteristic VOCs from the microbes [121]. Consequently, the detection of these VOCs may serve as a potential means for rapid and non-invasive microbial identification. The urea breath test for *H Pylori* detection and hydrogen breath test for gastrointestinal bacterial overgrowth are existing examples of non-invasive clinical breath tests [122]. Future studies may therefore seek to characterise the role of the cancer associated microbiome in the production of VOCs of interest.

There have been a number of studies investigating the headspace above bacterial cultures [123].Using PTR-MS, Crespo *et al*, have measured the differences between the headspace of mycobacterium kansasii and mycobacterium avium detecting dimethyl sulfide, dimethyl disulphide and propanethiol to be higher for Kansaii species.

The gastric microbiota has been studied using acid-reducing drugs in subjects with hypochlorhydric conditions relying on cultivation approaches [13,124]. These culture-based studies may have a skewed representation since a large portion of the microbes inhabiting in the gut have not yet been cultivated[125]. Although there are several molecular tools to characterize these microbes, very few studies have used culturing as a method to investigate the microbial composition of the stomach and oesophagus and their ability to produce VOCs of interest that can be used as potential diagnostic markers.

The microbial composition in oesophago-gastric cancer that is described in the previous chapter and published literature [48,49,50] has demonstrated an increase in Firmicutes phyla. These include microbial species such as Lactobacillus fermentum, E. coli, Streptococcus salivarius, Streptococcus mitis, Streptococcus anginosus, Klebsiella pneumonia and Clostridium. This chapter utilises in-vitro and ex-vivo culturing techniques in order to analyse the VOCs within the headspace of selected cancer associated bacterial species using GC-MS.

7.2 In-vitro culturing

7.2.1 Aims

i. The aim of this work is to define a link between oesophago-gastric cancer associated bacteria (Escherichia coli, Lactobacillus fermentum and Streptoccus salivarius) and target VOCs that are upregulated in oesophago-gastric cancer

7.2.2 Methods

7.2.2.1 *Media preparation and culturing bacteria*

The ability to differentiate between bacterial strains based on their VOC profile is greatest during the late exponential and static bacterial growth phases, when volatile emissions reach a maximum[126]. Consequently, the use of liquid medium is preferred for VOC sampling because the growth phase can be monitored more easily in liquid medium than in solid medium.

Headspace analysis were performed on bacteria sourced from commercial cell depositories (NCIMB), Escherichia coli (NCIMB9552), Lactobacillus fermentum (NCIMB11840) and Streptoccus salivarius (NCIMB701779). All bacterial strains were assessed for viability prior to conducting experiments. Preparation of the media was performed to enhance the growth of each bacterium using the appropriate pH adjustments. Nutrient agar, MRS broth media and Yeast glucose agar were manually prepared and used to culture Escherichia coli, Lactobacillus fermentum and Streptoccus salivarius respectively, with all constituents included in Table 20.

MRS Broth (Oxoid CM 359)	**Amount**	**Nutrient Agar**	**Amount**	**Yeast Glucose Agar**	**Amount**
Peptone	10.0g	Lab-lemco/beef extract	1.0g	Glucose	20.0g
Lab-lemco powder	8.0g	Yeast extract	2.0g	Yeast extract	10.0g
Yeast extract	4.0g	Peptone	5.0g	Agar	15.0g
Glucose	20.0g	NACL	5.0g	Distilled water	1.0L
Sorbitonmono-oleate	1.0mL	Agar	15.0g		
K2 HPO4	2.0g	Distilled water	1.0L		
Sodium acetate H_20	5.0g				
Triamonim citrate	2.0g				
$MgSO_4$ H_20	0.2g				
$MnSO_4$ H_20	0.05g				
Distilled water	1.0L				

Table 20: Constituents of MRS broth media, Nutrient agar and Yeast glucose agar used to culture Escherichia coli, Lactobacillus fermentum and Streptoccus salivarius respectively.

Dried cells from all three lyophilised strains (Escherichia coli, Lactobacillus fermentum and Streptoccus salivarius) were recovered and each single colony was placed on a plate and incubated at 37^{o}C within its appropriate media. Test results had shown positive growth for all strains, in addition to conducting cell banks after measuring their optical density using the spectrophotometer.

The following day, 10ul of bacterial cells and 2mL of fresh media were placed in sterile GC vials (Supelco 27136; Merck KGaA, Darmstadt, Germany) with crimpled lids. Vials were sparged with Argon before and after the addition of bacteria in order to facilitate anaerobic growth. Vials were incubated at 37 °C and three replicates of each examined strain were used to ensure data repeatability.

For running large-scale sampling analysis, it was necessary to use sufficient controls to verify sterility and define volatile background levels without the inoculum. Blank samples with uninoculated media as well as empty vials were therefore analysed. Figure 32 below shows the cultured bacteria placed in vials.

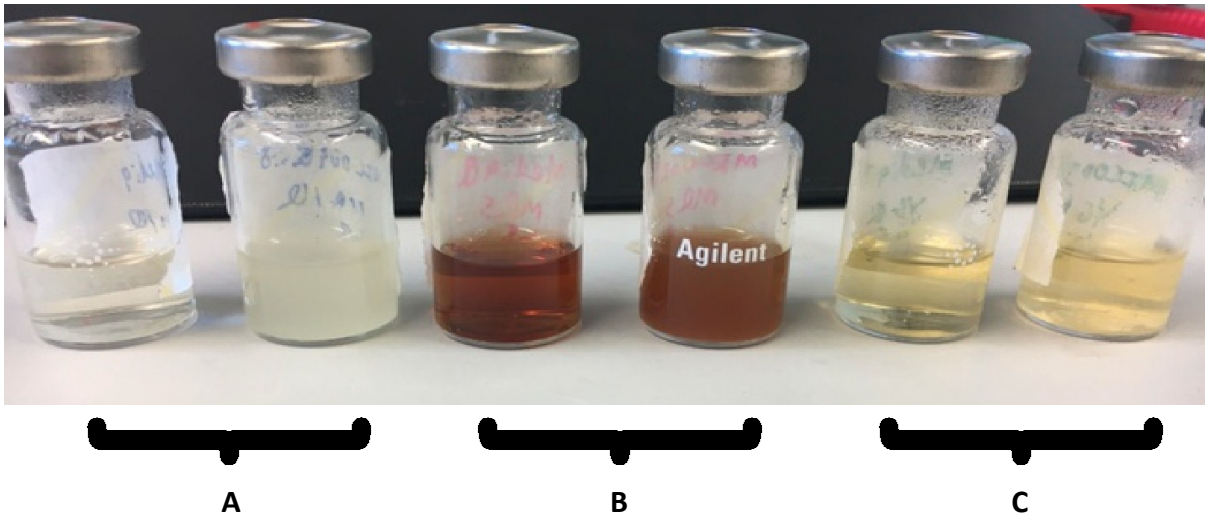

Figure 32:GC vials containing each of the three bacteria cultured. A (E coli), B (L fermentum) and C (S salivarius). Left hand vials contained only the media whilst the right side vials contained bacterial strains.

7.2.2.2 *VOC collection from strains*

Volatiles were collected in the absence of airflow, using a SPME technique. This technique utilises a fused-silica fibre coated with an appropriate stationary phase housed within a protective needle [127].As compared with dynamic headspace sampling, SPME has been extensively used in headspace analysis of bacterial cultures [127]. SPME is a static headspace sampling method that does not require air circulation for concentrating volatiles since it involves the use of a fibre coating to capture VOCs using a predetermined extraction time and incubation temperature. For a detailed overview of SPME optimisation parameters, the protocol by Risticevic *et al.* was consulted [128].The SPME fibre used for static headspace sampling was Divinylbenzene/Carboxen/Polydimethylsiloxane (Supelco 57329-U; Merck KGaA). Prior to usage, SPME fibre was pre-conditioned according to manufacturers guidelines.

7.2.2.3 *Sample analysis by GC-MS*

Headspace VOCs of cultured samples was analysed using a Perkin Elmer Clarus 580 GC with SQ8S MSD (PerkinElmer LAS Ltd, Bucks, United Kingdom) at 24 and 48 hours. Desorption of VOCs was performed by exposing the fibre coating in the GC inlet at 250 °C for 5 min. The VOC mixture was separated on a DB-624 capillary column (20 m × 0.18 mm ID × 1.00 µm d*f*; Agilent Technologies, Cheshire, UK) programmed at 0.5 mL/min constant Helium carrier. Oven temperature was set at 40 °C initially for 2 min, ramp to 240 °C at 20 °C/min with 3 min hold. The MS transfer line was maintained at 260 °C with 70 eV electron impact at 180 °C was set. MS analyser was set to acquire over the range of 25 to 250 m/z with data

acquisition approximated to 0.2 scan/sec. The benefits of this method of extraction for VOC analysis includes the limited amount of sample preparation steps required, lack of solvent use, and the potential to concentrate analytes and improve detection limits.

Control checks were regularly carried out to confirm the sensitivity/background levels in the GC–MS instrument. The use of a GC column was dedicated to VOC analysis and installation of a new liner and septum was considered. The MS instrument was also tuned (i.e., autotuned) and volatile standard mixtures were run as a quality control (QC) reference before starting experiments. Spectra for the QC standard were reproducible, including the peak widths, areas, retention times, and chromatographic resolutions for the volatiles. In order to determine whether changes in sensitivity or resolution were due to SPME performance or GC–MS detection level, QC VOC referencing was performed using liquid injection and results compared with previous ones. All samples were analysed as technical triplicates.

Volatile Organic Compounds released from the headspace of bacterial strains were selected based on previous VOC studies in oesophago-gastric cancer. These include acetone, acetic acid, butyric acid, pentanoic acid, hexanoic acid, phenol and Acetaldehyde. Calibration curves were produced using standard solutions of these compounds.

7.2.2.4 *Data analysis*

Data acquired from PerkinElmer GC-MSD system was processed using MzMine 2.30 [129], subjected to MetaboAnalyst 4.0 [130] for statistical analysis. Compound were identified using retention indices of authentic standards and matching of results with the NIST Mass Spectral Library version 2.2.

7.2.3 Results

In vitro headspace analysis by GC-MS demonstrated differences in the mass spectral profiles from each microbial culture. A representative chromatogram from each of the three strains and un-inoculated medium is provided in Figures 33,34 and 35 below, these show both the bioactive VOCs that are eluted and the background of VOC released from the medium.

For the first 24 hours of growth, little changes in the VOCs emitted from cultures and growth medium was observed for all the bacteria. After 48 hours, the levels of Phenol and Acetic acid increased from the headspace of Escherichia Coli strain and not from the medium. The L fermentum strain showed a similar trend with increased acetaldehyde, ethanol and 2–pentene levels detected above cultures. S salivarius, produced increased levels of ethanol and 2-Pentene.

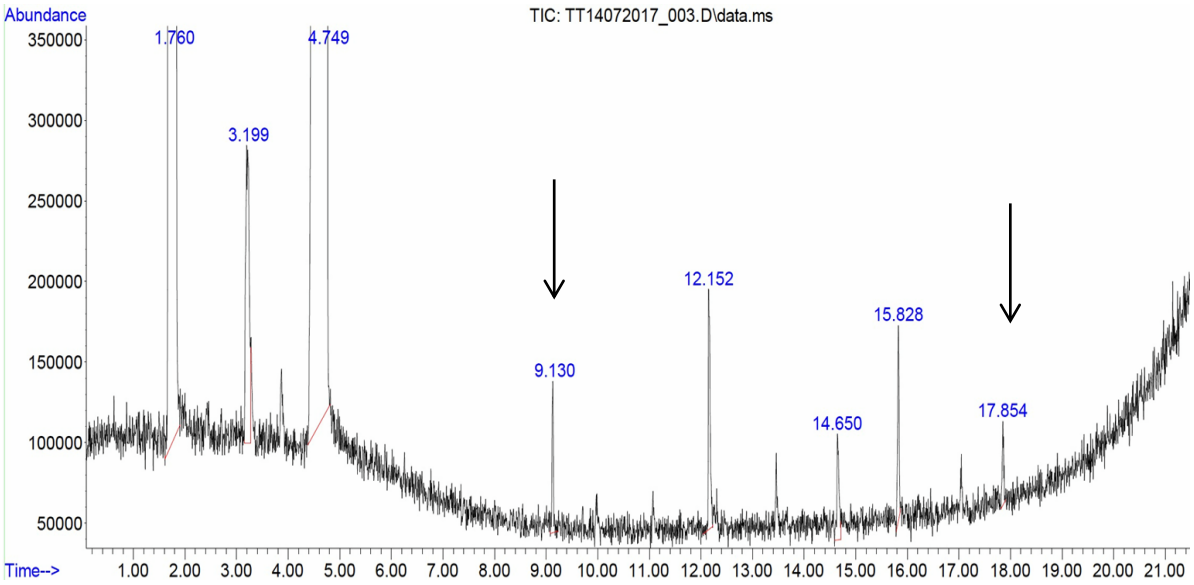

Figure 33:Chromatogram showing acetic acid (retention time 9.1) and phenol (17.8) peaks produced by E coli.

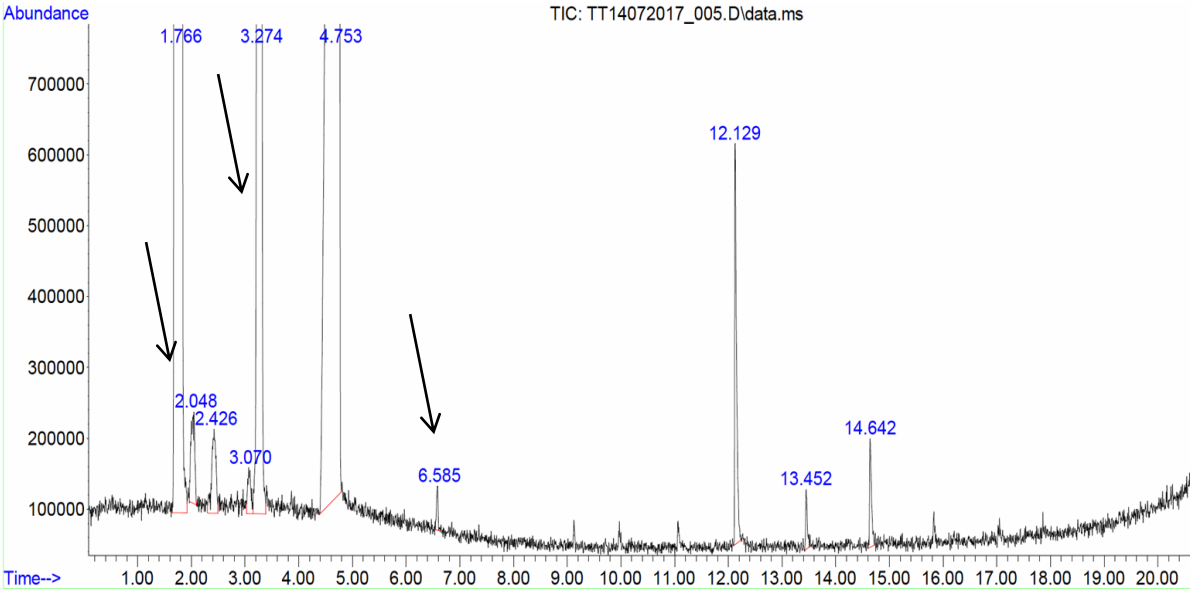

Figure 34:Chromatogram showing acetaldehyde (retention time 2.048), ethanol (3.2) and 2-pentene (6.5) peaks produced by L fermentum.

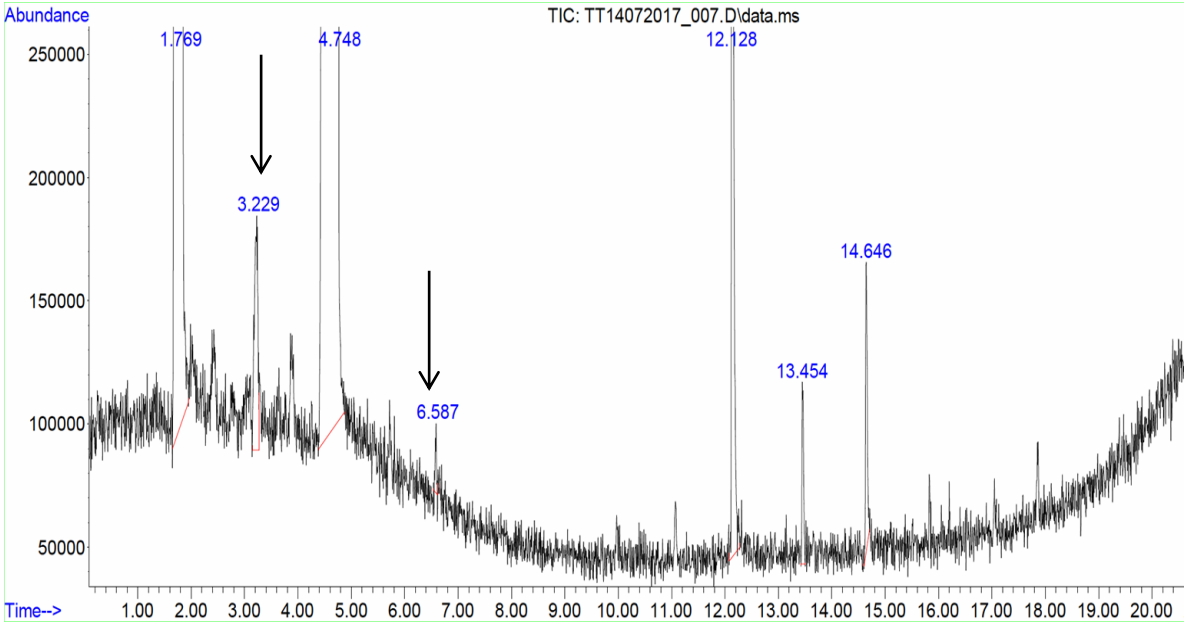

Figure 35:Chromatogram showing ethanol (retention time 3.2) and 2-pentene (6.5) peaks produced by S salivarius.

7.2.4 Discussion

Volatile compounds are known to be emitted by microbes and have significant roles in modulating inter- and intra-cellular communications [131.] Although solvent extraction of biological materials is adequate for obtaining a considerable fraction of non-volatile metabolites present within the bacteria, headspace analysis provides a more representative and convenient method of sampling VOCs. Given that various microbial cell cultures emit characteristic aromas, which can be precisely determined by different mass spectrometry methods to provide a foundation for assignment and discrimination[132].

This study demonstrated that various bacteria found to be upregulated in oesophago-gastric cancer have the ability to produce VOCs that are significantly different in exhaled breath of cancer patients compared to healthy controls. Out of the principal compounds that have previously been reported to be upregulated in the exhaled breath of patients with oesophago-gastric cancer, phenol and acetic acid were produced by headspace of E coli while acetaldehyde by L fermentum.

Acetic acid, which is a metabolic intermediate for the generation of acetyl-CoA. In our previous study analysing the headspace of non-derivatised tissue samples, we observed higher concentrations of acetic acid and phenol in both in the cancer cohort compared to healthy controls. Zhang *et al.* performed NMR spectroscopy of blood samples from patients with oesophageal adenocarcinoma and reported changes in the trichloroacetic acid cycle were dominant factors in the biochemistry of this cancer[133]. The hypothesis of bacterial conversion of tyrosine to phenol within the upper gastrointestinal tract using tyrosine provides a potential explanation for their increased concentrations[134]. For acetaldehyde,

studies have revealed that microbes are responsible for the bulk of its production from ethanol by possessing alcohol dehydrogenase (ADH) activity both in saliva and *H pylori*-infected and achlorohydric stomach[135].

The protocol for analysis of VOCs described herein was optimised for targeted *in vitro* analysis of volatiles in microbial samples. Absolute quantification of volatiles from bacteria using a static SPME technique is challenging, considering the differences in volatile extraction depending on incubation time, different fibre type or response. Consequently, this selection can limit sensitivity by preferentially adsorbing or excluding VOCs based on polarity or size [136].

The number of VOCs that can be detected identified, and/or quantified using the current method was another challenge of this study. A time-of-flight MS (TOF/MS) detector for VOC analysis, which was not available in our laboratory at that time this experiment was conducted, would have been a better analytical platform with a higher sensitivity level, especially when coupled to 2D column gas chromatography. This technology offers several advantages, including notably fast scan times that give rise to either improved peak deconvolution or reduced run times which could also have helped in determining the chemical structure of unknown volatile peaks.

These findings support the data of the previous work and prove that bacteria can potentially produce similar compounds to what has been present in headspace analysis of tissue, exhaled breath, gastric content and urine of patients with oesophago-gastric cancer. Bacterial VOC could however co-elute leading to overlapping MS spectra, which hinders the process of identifying specific volatiles [137]. However, the data presented suggests that the

use of VOC emissions as a signature of specific bacteria is plausible and it might ultimately be achievable through breath analysis, although much more research will be needed if this is to become a viable proposition.

7.2.5 Conclusion

This experiment has demonstrated the potential of VOC production from upregulated bacteria present in oesophago-gastric cancer with the use of in vitro measurements. Real-time non-invasive diagnosis of bacterial presence in the body could enable us to link the microbiome present in oesophago-gastric cancer and VOCs present in breath testing.

7.3 Ex-vivo culturing

7.3.1 Aims

The aim of this study is to perform ex-vivo culturing work to isolate the predominant microbial species found within oesophago-gastric adenocarcinoma tissue samples and to detect their potential role in the production of VOCs.

7.3.2 Methods

7.3.2.1 *Patient selection*

Cancer samples were collected from *H Pylori* negative patients undergoing surgical resection of oesophago-gastric adenocarcinoma (n=8). Matched healthy gastric (n=8) and oesophageal (n=6) mucosa samples from the same cancer patients were also collected. Patients with a normal upper gastrointestinal tract on OGD (n=8) were recruited as 'healthy' controls. Following collection all samples were placed in vials containing freeze media and snap frozen in liquid nitrogen before being stored at -80°C prior to analysis.

7.3.2.2 *Culturing bacteria*

Tissue samples were defrosted and media was removed by centrifugation and washing with PBS. Each tissue sample was then placed in in a 4 mL headspace vial (Supelco 27136; Merck KGaA, Darmstadt, Germany) with sterile phosphate buffer solution and lysed using a medium speed vortex. Supernatant was placed into prepared FAA plates (Sigma-Aldrich; Merck KGaA, Darmstadt, Germany) and incubated at 37°C overnight.

After incubation for 48hrs the predominant bacterial colonies have accumulated to form a visible patch on the plates, which were picked and categorized according to morphological features including shape, appearance, size and texture (Figure 37). The colonies were then placed into 100ul PBS. The remaining bacteria present on the cultured plate was then wiped out and also suspended into 100ul PBS after centrifuging. Figure 36 below illustrates the culturing steps on both days prior to GC-MS analysis.

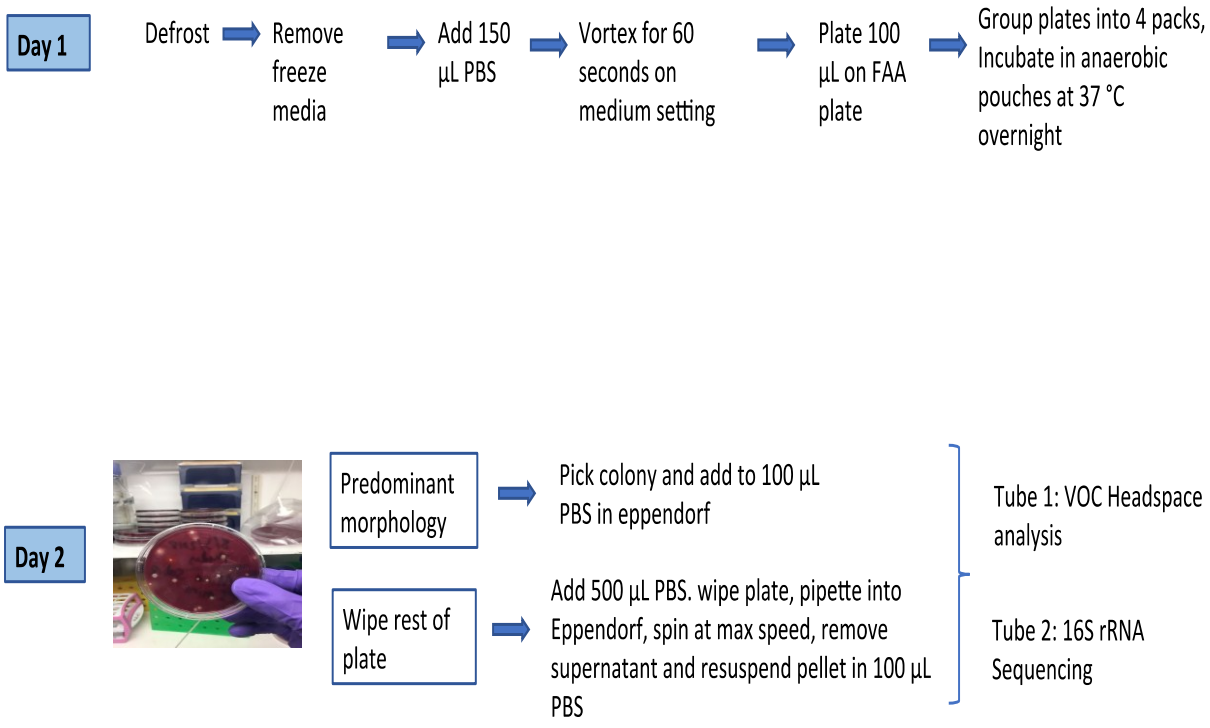

Figure 36:Steps for bacterial culturing, headspace analysis and 16S rRNA sequencing.

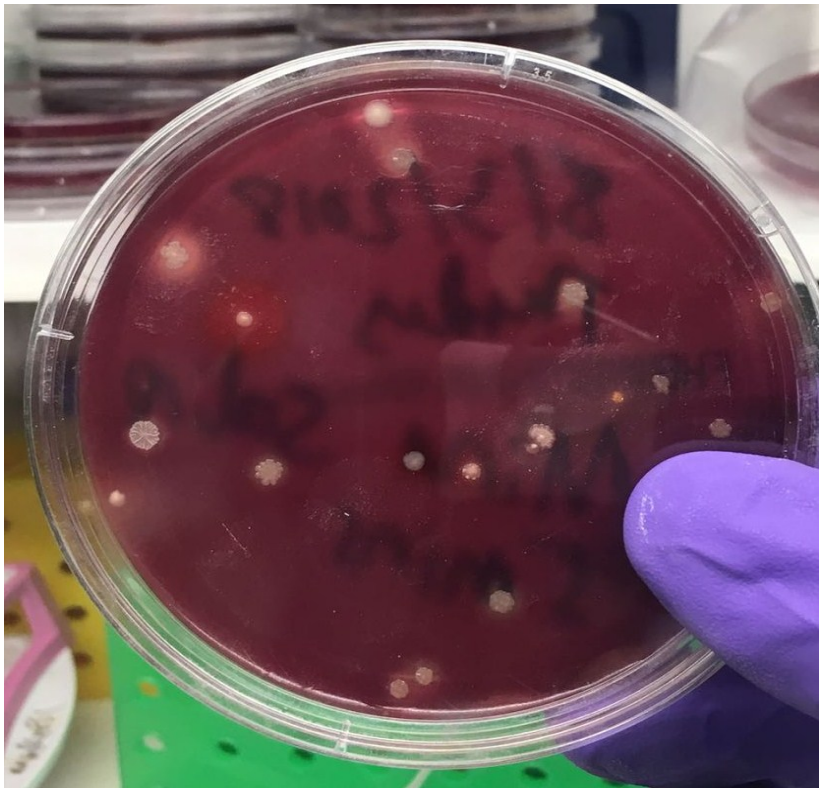

Figure 37:An example of bacteria grown on a plate demonstrating few colonies to be similar in shape and appearance.

7.3.2.3 *GC-MS analysis*

Headspace VOC of cultured samples was analysed using a Perkin Elmer Clarus 580 GC with SQ8S MSD (PerkinElmer LAS Ltd, Bucks, United Kingdom). Blank samples with uninoculated media were used as controls to eliminate any VOCs released from growth medium. Bacterial late exponential and static growth stage was initially determined to allow volatile emissions reach maximum. The MS instrument was tuned and volatile standards were tested as a quality control reference before starting experiments. Headspace analysis was undertaken via solid phase micro-extraction (SPME) and direct liquid injection to monitor real-time variations in compounds.

SPME was performed using a carboxen/polydimethysiloxane SPME fibre in 20mL screw-top glass vials with a PTFE septum. Prior to usage, SPME fibre was pre-conditioned according to the manufacturers guidelines. SPME extraction was performed at 60°C with intermittent agitation at 500rpm. Volatiles were collected in the absence of airflow, after 24 and 48 hours of incubation followed by direct release into a heated gas chromatography injector.

Direct liquid injection sample analysis was achieved by injecting 1 µL of sample aliquot from the broth culture using a PALcombi-XT autosampler (CTC Analytics AG, Switzerland) into the PerkinElmer GC-MSD inlet at 250 °C. VOCs were separated on a Stabilwax-DA capillary column (30 m × 0.25 mm ID × 0.25 µm d*f*; Restek Corporation, PA) programmed at 1.0 mL/min constant Helium carrier. Oven temperature was set at 40 °C initially for 5 min, ramp to 240 °C at 10 °C/min with 5 min hold. The MS transfer line was maintained at 260 °C with 70 eV electron impact at 180 °C was set. MS analyser was set to acquire over the range of 40 to 400 m/z with data acquisition approximated to 0.5 scan/sec.

7.3.2.4 *GC-MS data analysis*

Data acquired from PerkinElmer GC-MSD system was processed using MzMine [129], 2.30 subjected to MetaboAnalyst 4.0 [130] for statistical analysis. Compound were identified using retention indices of authentic standards and matching of results with the NIST Mass Spectral Library version 2.2.

7.3.2.5 *16S rRNA sequencing and analysis*

Using the method and analytical strategy described in the previous chapter, DNA extraction and 16S rRNA sequencing was performed on all of the thirty isolated bacterial samples using the Illumina MiSeq platform and Mothur software for analysis, aiming to detect the identity of the predominant bacterial species found in tissue samples.

7.3.3 Results

Samples were collected from 8 oesophago-gastric adenocarcinoma patients and 8 controls with normal upper gastrointestinal tract. Of the gastric cancer cases the majority of cases had advanced disease with 7/8 having T-Stage T3/T4 disease and 6/8 having lymph node involvement. Demographic details for each group are shown below in table 21.

	Control (n=8)	**Cancer (n=8)**
Age (Median)	55.5 (35-86)	65(34-78)
Male	4 (50 %)	6 (75 %)
PPI Use	4 (50 %)	7 (88%)
Aspirin use	0	1 (13%)
Hypertension	1 (13%)	4 (50 %)
Diabetes	2 (25%)	3 (38%)
Statin Use	1 (13%)	1 (13%)
Non-Smoker	3 (38%)	4 (50 %)
Alcohol <21 units/week	5 (63%)	6 (75 %)

Table 21:Patient details showing the demographic differences between the control and cancer patient groups.

7.3.3.1 *VOC analysis*

From the compounds of interest, headspace and direct liquid injection analysis of predominant bacteria present on the plates demonstrated acetic acid and butanoic acid to be significantly higher in samples derived from cancer patients compared to healthy controls. No significant differences in VOC levels were observed between bacteria present in cancer patients and matched normal stomach and oesophagus from the same patient. Bacteria present in healthy control samples produced higher levels of benzaldehyde compared to cancer samples.

The software algorithm detected the fragment ions of acetic acid and benzaldehyde repetitively, which is likely due to splitting of the peaks at almost similar retention times. This could also be the detection of isotopic ions of both compounds (acetic acid, m/z 61 and 62, benzaldehyde, m/z 77 and 106). Compounds that were significantly higher in bacteria present in cancer versus controls are shown in table 22 and Figures 38,39 and 40 below.

Name of compound	m/z	Fold change	P value
Acetic acid	3/62.0mz/7.571min	22.17	0.017585
Acetic acid	118/61.0mz/7.490min	8.3056	0.018511
Benzaldehyde	39/77.1mz/7.858min	0.39239	0.017585
Benzaldehyde	28/106.1mz/7.859min	0.45503	0.018511
Benzaldehyde	40/77.1mz/7.859min	0.40552	0.039901
Butanoic acid	79/60.0mz/9.254min	3.8358	0.031006
Propyl acid	83/92.1mz/9.270min	0.36372	0.042575
Phenylethanal	74/92.1mz/9.271min	0.33701	0.017585

Table 22:Relative fold change of compounds produced by the predominant bacteria showing significant compounds at higher concentrations in cancer versus control patients.

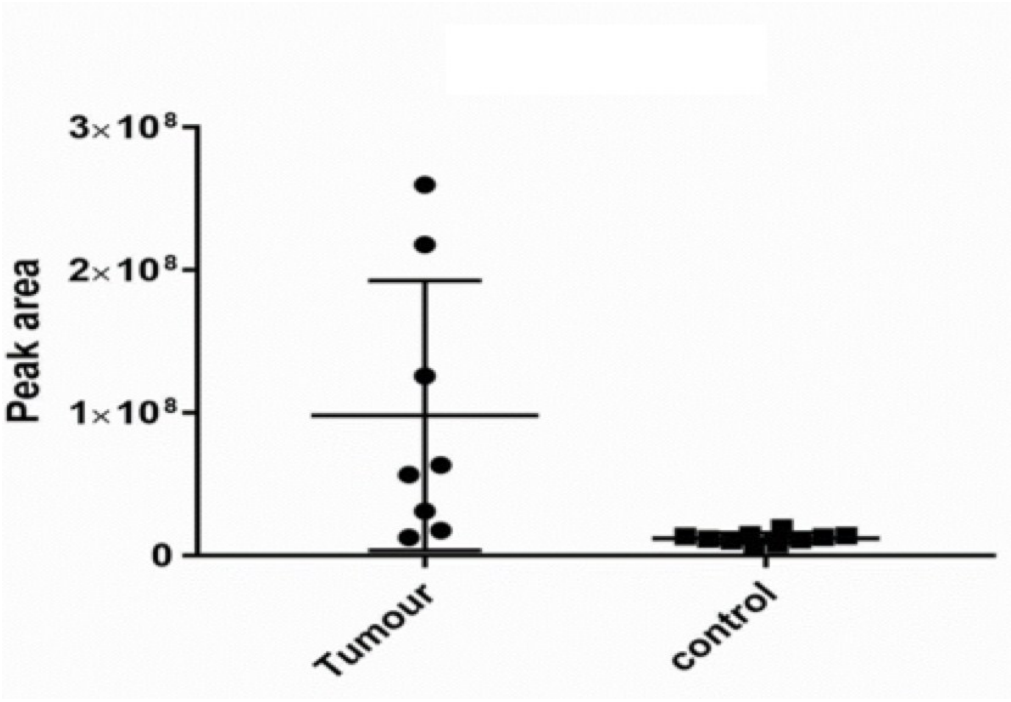

Figure 38:Scatterplot for all the predominant bacteria present on plates showing acetic acid (p=0.0175) to be higher in cancer versus control patients.

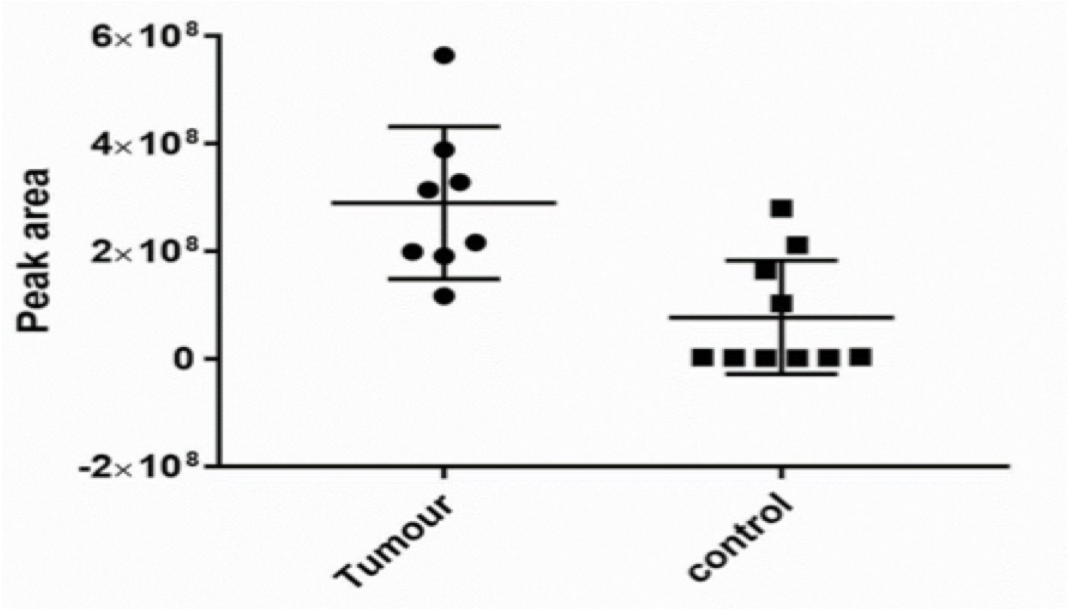

Figure 39:Scatterplots for all the predominant bacteria present on plates showing butanoic acid (p=0.031) to be higher in cancer versus control patients.

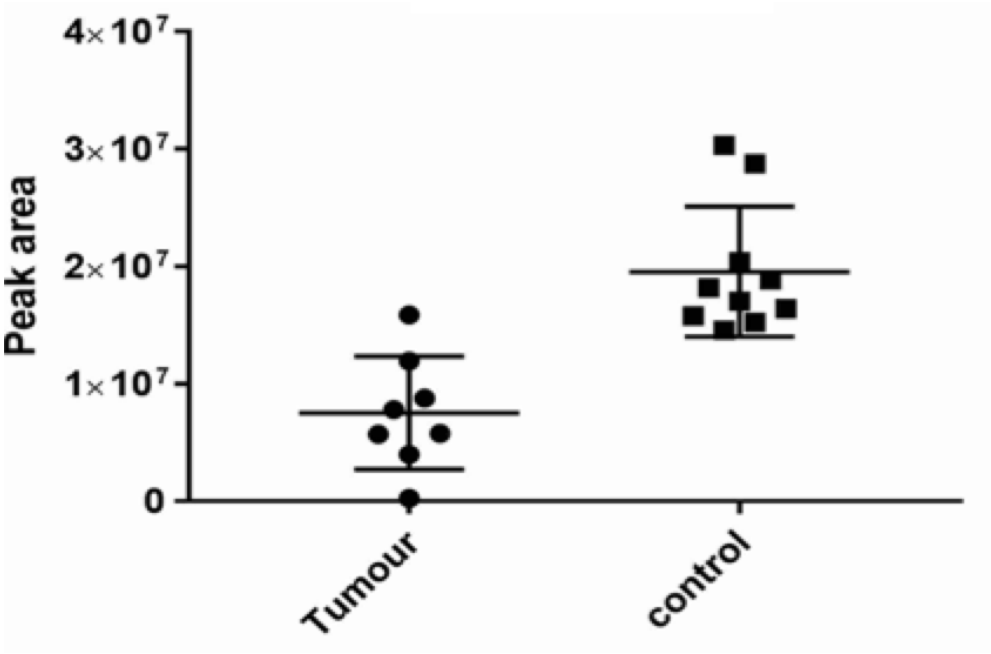

Figure 40:Scatterplots for all the predominant bacteria present on plates showing benzaldehyde (p=0.039) to be higher in control patients more than cancer.

7.3.3.2 *16S rRNA analysis*

Each of the eight control samples demonstrated consistency in the prevalence of Streptococcus dentisani species belonging to the Firmicutes family. Streptococcus anginosus known to be from the Firmicutes family was present in three tumour samples. Actinomyces odontolyticus, Prevotella plallens, Veillonella dispar and Streptococcus salivarius, belonging to Actinbobacteria, Bacteroidetes and Firmicutes families respectively, were each present in only one tumour sample as shown in Figure 41.

Streptococcus salivarius and Streptococcus anginosus were both found in the previous chapter at higher concentration in gastric cancer versus control patients, and have found to be producing higher concentrations of acetic and butanoic acid using GC headspace and liquid injection analysis.

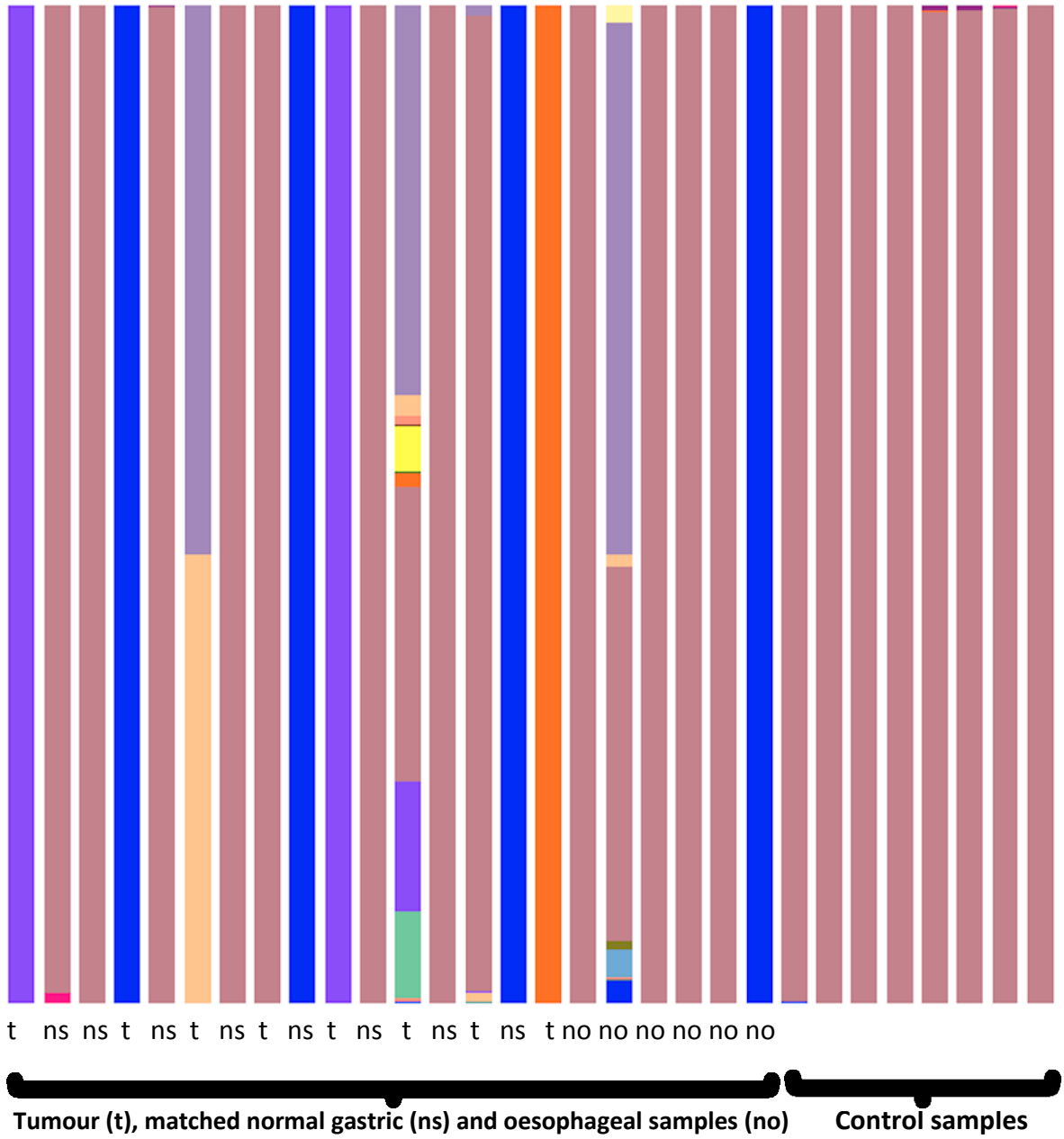

Figure 41:Illustrating the differences in bacteria between cancer and control tissue samples. Streptococcus dentisani (pink colour) was only present in samples from control patients and in matched healthy stomach (ns) and oesophagus (no) from the same cancer patient. As for tumour samples(t), each one demonstrated the presence of Streptococcus anginosus (Purple), Actinomyces odontolyticus (Blue), Veillonella dispar (light purple), Prevotella plallens (yellow) and Streptococcus salivarius (orange).

7.3.4 Discussion

Ex-vivo headspace analysis by GC-MS has shown that there are sufficient differences in the mass spectral profiles from different microbial cultures to permit rapid discrimination. This study analysed headspace and direct liquid injection of abundant microbial phyla from tumour biopsies of patients with oesophago-gastric cancer and demonstrated how they can potentially play an important role in tumorigenesis.

Tissue samples from healthy control patients have all shown a similar trend of having Streptococcus dentisani species belonging to the Firmicutes family in all of them. Also from the Firmicutes family, Streptococcus anginosus and Streptococcus salivarius were found to be present in tumour samples producing the highest concentrations of acetic and butanoic acid compared to the remaining tumour samples. These bacteria correlate with the data demonstrated in the previous chapter confirming their presence in oesophago-gastric cancer patients.

In this study acetic acid was present at higher concentrations in bacteria from cancer versus controls samples. In the previous study (section 2.2) the headspace of non-derivatised tissue samples was found to contain higher concentrations of acetic acid in the cancer cohort compared to healthy controls. Previous studies have reported that the aliphatic fatty acids, butanoic acid, pentanoic acid, are important in regard to tumourigenesis and response to anti cancer therapy [87,138,139,140,141]. These fatty acids may be ingested as part of a normal diet but are thought to be produced primarily by the gut microbiota during the fermentation of dietary fibre[87].

It should be noted that bacteria produce and emit exceptionally diverse organic and inorganic volatile compounds. Inorganic volatile compounds such as carbon dioxide (CO2)

hydrogen cyanide (HCN), hydrogen sulfide (H2S), ammonia, and nitric oxide (NO), are also liable to accumulate within the closed volatile collection setup and potentially may have an effect on growth promotion [143]. Although the results from this study are encouraging for the use of headspace and direct analysis above microbial cultures, it remains unclear whether current mass spectrometric analytical techniques have sufficient sensitivity to identify specific bacteria by analysing the VOC signature within exhaled breath. Another important issue to address is the definitive identification of the VOCs emitted from microbial cultures, for which there has only been a limited amount of work to date. Such assignments would shed light on the metabolic processes occurring in the cultures, which are of considerable interest to microbiologists and may aid in the development of new techniques.

7.3.5 Conclusion

VOCs detected within bacterial headspace are similar to compounds that have been found in tissue biopsies, biofluids and exhaled breath from cancer patients across several mass spectrometry platforms. The work has developed reliable methods for the high throughput analysis of VOCs utilising a combination of complementary mass-spectrometry based platform for gas phase analysis. Whilst the mechanisms underlying these changes remain incompletely understood, they are expected to involve important intracellular pathways related to cell metabolism and tumorigenesis.

These finding offer new impetus for future investigations that should seek to define not only the endogenous origin of target VOCs but also their place with the care pathway of oesophago-gastric cancer patients.

8 AUGMENTED MICROBIOME BREATH TEST FOR OESOPHAGO-GASTRIC CANCER (AMBEC)

8.1 Background

Based on the detection of small molecules in exhaled breath, our research group has developed a non-invasive test for detection of oesophago-gastric cancer. Results presented in previous chapters indicate that the tumour associated microbiome could play an important role in the production of VOCs in oesophago-gastric cancer. Experiments identified a unique population of bacteria that were associated with oesophagogastic tumours including: Lactobacillaceae, Streptococcaceae and Eubacteriaceae that are from the Firmicutes phyla. Isolated cultures of these bacteria were found to emit the same VOCs that were found to be elevated in the breath, urine, gastric content and luminal air of patients with oesophago-gastric cancer [19-23].

The breath test for oesophago-gastric cancer diagnosis initiated from volatile biomarker discovery to a multicentre NIHR-funded clinical study that demonstrated 80% sensitivity and 81% specificity[19]. Rather than engage in a law-of-diminishing-returns-type attempt to increase the sensitivity of test-equipment, it may be possible to augment the diagnostic response using simple nutrients. It is hypothesised that cancer cell and associated bacteria will be able to utilise administered substrates within defined metabolic pathways responsible for the production of VOCs. By exploiting deregulated metabolic pathways in this way, we expect to observe a transient elevation in cancer associated VOCs, thus increasing the accuracy of the test.

The augmented microbiome mediated breath test (AMBEC) will help in addressing the current diagnostic dilemma for oesophago-gastric cancer. Red flag symptoms often indicate a late-stage cancer with poor prognosis whilst, at the same time, there is a high prevalence of non-specific symptoms relative to cancer incidence. With ready access to an AMBEC, a GP would not need to watch-and-wait to see if symptoms worsen but could offer the primary care based test immediately in much the same manner as a routine blood test. A nurse would perform the test, sending breath samples to a regional laboratory for analysis. A positive result would warrant immediate referral for endoscopy. A negative result would permit the GP to reassure the patient and offer retesting if symptoms persist.

The following chapter is divided into three section: (i) pilot study to assess the viability of a stimulant mix that contains specific nutrients, on limited number of bacteria present in oesophago-gastric cancer and their VOC response; (ii) high throughput study to maximise VOCs response by optimising the stimulant mix and testing the response on an increased number of in-vitro targeted bacteria using a more sensitive mass spectrometry technique, and; (iii) A clinical study to examine the augmented VOC response of cancer and non-cancer patients to an oral stimulant drink containing the same substrates used in the stimulant mix given to targeted bacteria in the high throughput study.

8.2 **Pilot study**

8.2.1 Background

This experiment aimed to demonstrate whether a preliminary stimulant mix can stimulate VOC production by selected oesophago-gastric cancer associated bacteria. A Pilot study was conducted for *Escherichia coli* (NCIMB 9552) and *Lactobacillus fermentum* (NCIMB 11840), assayed under micro-aerobic conditions after a brief period of starvation.

8.2.2 Aims

i. To define the preliminary components of the proposed oral stimulant drink.

ii. To determine the VOC response of selected tumour associated bacteria to a defined nutrient challenge.

8.2.3 Methods

8.2.3.1 *Preparation of the proposed oral stimulant drink*

Preparation of the stimulant mix was performed in collaboration with Ingenza Ltd. Substrates were selected based on their established association with bacterial metabolism. These substrates were derived from: (i) the dataset of gastric bacterial isolates most commonly associated with cancer tissue; (ii) extensive bioinformatic review of the enzymatic pathway regulation and biochemical flux in key bacterial species (Table 23), and; (iii) scientific literature describing the conversion of particular primary metabolites to specific VOC species. Sugars, organic acids and amino acids were identified as priority

compounds. Specific substrates selected for the initial stimulant mix are presented in Table 24. Several fatty acid substrates were also considered appropriate but excluded due to insolubility within the aqueous formulation.

A stock solution of tyrosine (10.56 mM) was prepared volumetrically in aqueous hydrochloric acid (50 mM). The acidic mixture was stirred vigorously with heating until fully dissolved. A mixture of the remaining stimulants was prepared in a separate volumetric flask. A sufficient volume of the tyrosine stock was added to the stimulant mixture to give a final concentration of 1 mM when prepared volumetrically in PBS (0.01 M). If required, the solution was adjusted to pH 5 using dilute hydrochloric acid or sodium hydroxide. The final stimulant-PBS solution was filter-sterilised and stored at 4 °C until use.

Substrate	Associated Biomarker(s)	Scientific Rationale	Reference number
Tyrosine Phenylalanine	Phenols Cresol Ethylphenol	Research indicates that phenolic metabolites are primarily derived from the fermentation of tyrosine. Production of volatile phenols by Lactobacillus plantarum isolates	144 145 146
Glutamic acid	Phenol Cresol Ethylphenol	Glutamate is essential for transamination, allowing the production of aromatic amino acids via the Shikimate pathway.	147 148 149 150
Glucose	Acetic acid Acetaldehyde Phenol Cresol Ethylphenol	Glucose is fermented to a variety of organic acids under oxygen-limited conditions. These reactions ensure the replenishment of NAD^+ for continued glycolytic activity. Acetic acid fermentation is of particular benefit as it generates an additional molecule of ATP via substrate-level phosphorylation. Glycolysis also provides substrates for the Shikimate pathway, essential for the production of aromatic amino acids.	147 148 151
Xylose	Acetic acid Acetaldehyde Phenol Cresol Ethylphenol	Xylose is readily converted into xylulose, which enters the pentose phosphate pathway as xylulose-5-phosphate (D-xylulose-5-P). D-xylulose-5-P can subsequently enter glycolysis as glyceraldehyde-3-phosphate (GA3P) and progress towards pyruvate, both generating ATP and enabling mixed acid fermentation. D-xylulose-5-P can also be converted into erythrose-4-phosphate, an essential substrate for the formation of aromatic amino acids.	147 148
Lactose	Acetic acid Acetaldehyde Phenol Cresol Ethylphenol	Under glucose-starved conditions, lactose can be hydrolysed to liberate glucose and galactose as carbon sources.	147
Ethanol	Acetic acid Acetaldehyde	Alcohol/aldehyde dehydrogenases catalyse the oxidation of ethanol to acetic acid. These reactions replenish intracellular NADH for mixed-acids fermentation.	147 152
Glycerol	Acetic acid Acetaldehyde Phenol Cresol Ethylphenol 3-HPA	Glycerol is readily metabolised to GA3P, a key intermediate in glycolysis. GA3P can progress towards pyruvate, generating ATP and substrates towards organic acid fermentation and aromatic amino acid synthesis.	147 153
Stearic acid Oleic acid Palmitic acid	Acetic acid Butyric acid Hexanoic acid Acetaldehyde Hexanal	Long-chain fatty acids undergo oxidation to generate short-chain fatty acids and a pool of acetyl-CoA, which can be utilised for acetic acid fermentation under oxygen-limited conditions.	147

Table 23:showing the stimulant justification used to prepare the OSD drink.

Component	Pilot study mix
	g/L
Tyrosine	0.192
Phenylalanine	0.596
Glutamic Acid	0.469
Glucose	0.576
Xylose	0.458
Lactose	1.027
Ethanol	0.136
Glycerol	0.280
Sorbitol	0.552

Table 24:Summary of substrate masses used for the volumetric preparation of stimulant mix.

8.2.3.2 *Assay conditions*

Under aseptic conditions, a starter culture of *Escherichia coli* (NCIMB 9552) and *Lactobacillus fermentum* (NCIMB 11840) was grown aerobically in recommended media overnight (37 °C, 250 rpm). The culture was subsequently back-diluted to an OD_{600} = 0.1, incubated (37 °C, 250 rpm) and the cells were harvested during mid log phase (OD_{600} = 0.5 – 0.6). The biomass was pelleted by centrifugation (3000 rpm, 10 minutes) resuspended in PBS (0.01 M, pH 5, 50 mL) and starved for 1 hour under micro-aerobic conditions (37 °C, 250 rpm). The starved cells were equally divided into separate falcon tubes, pelleted by centrifugation (3000 rpm, 10 minutes) and resuspended in PBS solution (0.01 M, pH 5, 10 mL) with or without the stimulants. 50 mL of stimulant mix has been added to the cells. Stimulant-free and cell-free control experiments were also performed. All experiments were performed in duplicate.

Each experiment was subsequently incubated under micro-aerobic conditions (37 °C, 250 rpm) for 21 hours. Samples for analysis were acquired at t = 0.00, 0.25, 0.50, 0.75, 1.0 and 21 hours.

8.2.3.3 *GC-FID Analysis*

An Agilent 6890N gas chromatographic system (Agilent Technologies, Santa Clara, California, USA) equipped with a flame ionization detector (GC-FID) and an automated liquid sampler was used for method development and method validation. The GC-FID method has been developed for the detection of target phenols and aliphatic acids that are potential biomarkers for gastric cancer [154]. This method has since been applied to biotransformation samples to monitor the production of these biomarkers from a rich solution of aromatic

amino acids, saccharides and alcohols. The samples were centrifuged (14,800 rpm, 2 minutes) and 100-200 µL of the supernatants were isolated and filtered through a 0.2 micron syringe filter into a 2 mL sterile GC vials (Supelco 27136; Merck KGaA, Darmstadt, Germany) with crimpled lids. The vials were sealed with a crimp cap and queued on the GC autosampler carousel. GC-FID analysis was undertaken using the method parameters outlined in supplementary file 1.

8.2.4 Results

8.2.4.1 *E. coli (NCIMB 9552)*

In this experiment, the starved *E. coli* cells were resuspended in PBS solution containing the stimulant mix and GC-FID was employed to monitor the formation of potential biomarkers. Stimulant-free and cell-free control experiments were also performed. Acetic acid was the only biomarker that could be tentatively identified by GC-FID analysis of the stimulant-fed biotransformation (Figure 42); however, subtle differences in retention time and peak symmetry required further work to be done, including standard addition and mass characterisation which was needed to confirm the identity of the analyte. There are a number of possible metabolic pathways that could explain the production of acetate under anaerobic fermentative conditions from a sugar feedstock.

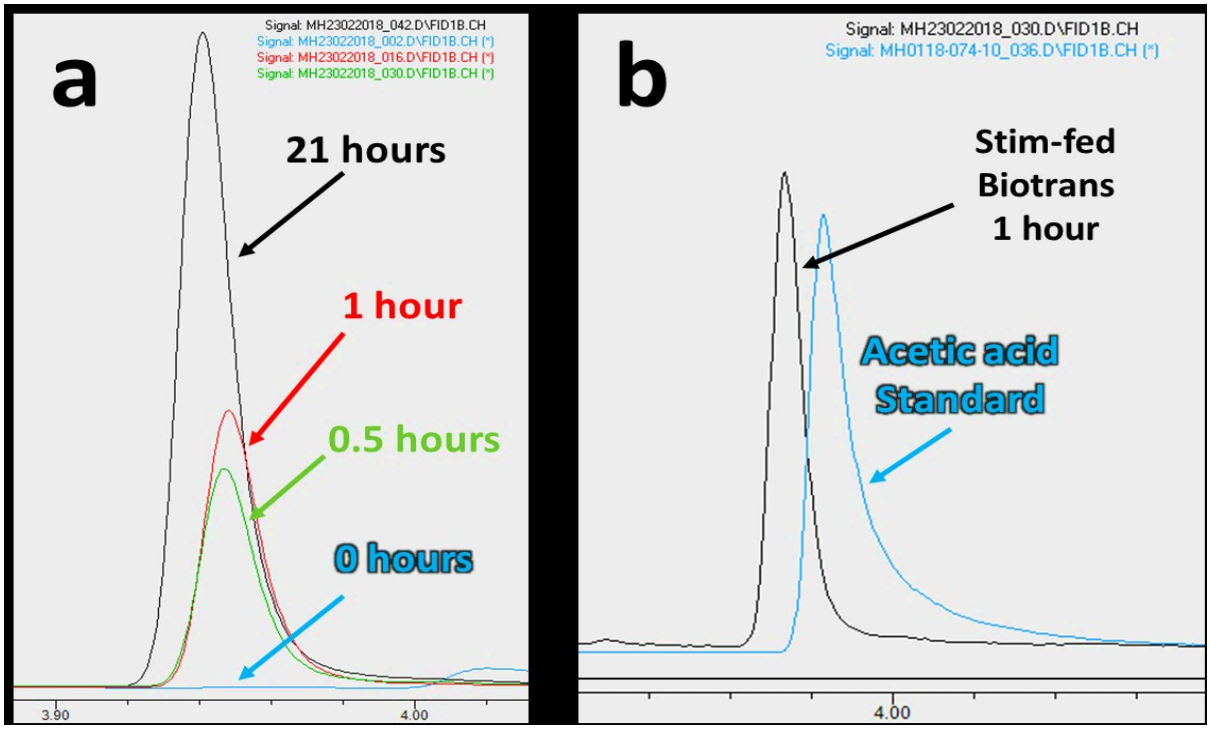

Figure 42: **a.** Formation of an analyte (retention time: 3.94 – 3.95 minutes) in the stimulant-fed biotransformation tentatively identified as acetic acid.

b. Comparison between an acetic acid reference standard and the observed GC-FID peak.

A background level of this analyte is also present in the cell-free control, however this did not impact the significance of the observed quantities of this compound produced in the stimulant-fed biotransformation (Figure 43).

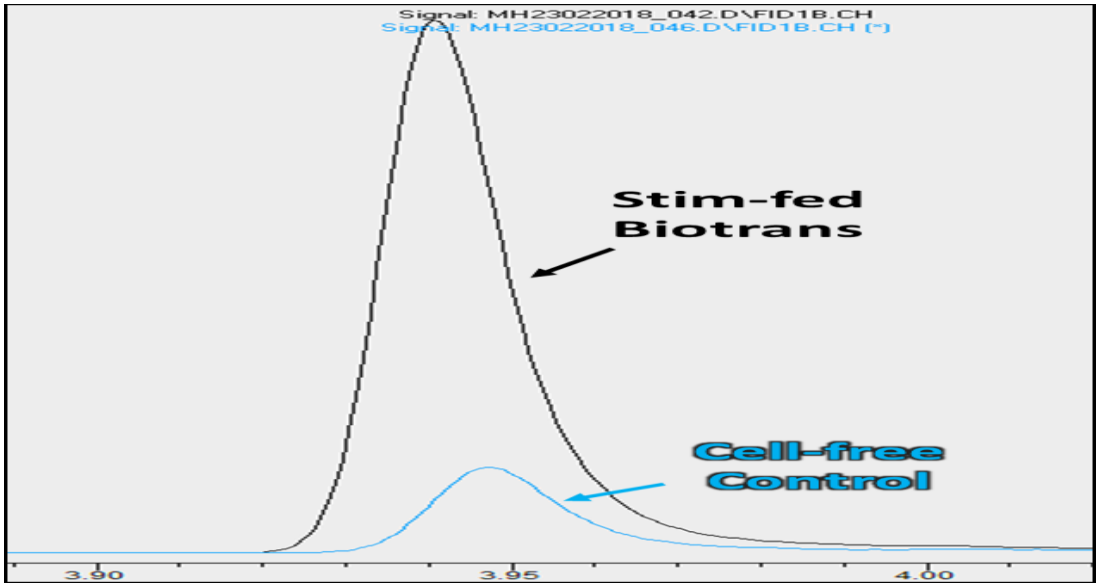

Figure 43:Peak comparison (retention time: 3.94 – 3.95 minutes) between the stimulant-fed biotransformation and cell-free control after 21 hours.

An additional unknown analyte peak was detected (retention time: 8.75 – 8.78 minutes) in the stimulant-fed biotransformation after 0.50 hours, which was absent in all stimulant-free and cell-free control samples (Figure44). This peak has been previously observed in a preliminary study on stimulant-fed *E. coli* (NCIMB 9552) under the same micro-aerobic conditions. Based on the retention time and peak symmetry, it is hypothesised that the compound in question is of mid-to-low polarity and possesses a boiling point within the range of 190 – 200 °C, however it was not possible to confirm the identity of the compound.

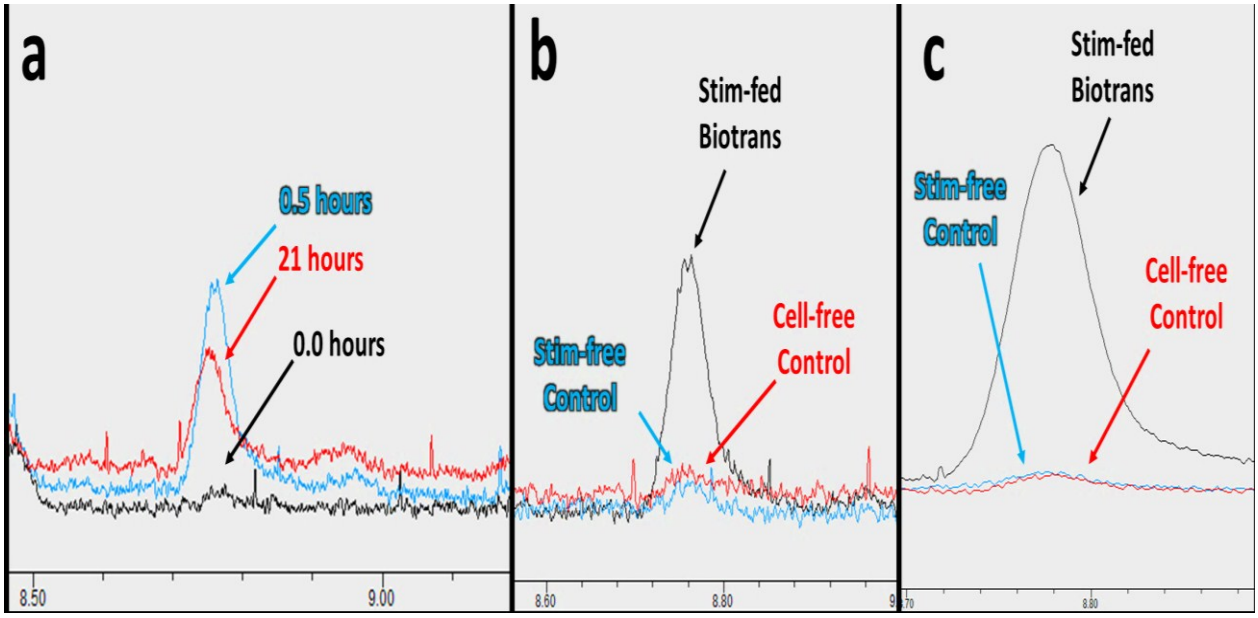

Figure 44: **a.** Formation on an unidentified analyte (retention time: 8.75 minutes) in the *E. coli* stimulant-fed biotransformation after 0.50 hours.

b. Chromatographic comparison between stimulant-fed, stimulant-free and cell-free experiments at t = 0.50 hours.

c. Chromatographic comparison obtained from an earlier preliminary study of *E. coli* under micro-aerobic conditions.

Another unique analyte was also observed in the stimulant-fed biotransformation after 21 hours. This peak is absent in both controls and is not observed in earlier time points (Figure 3.1.4). The retention time and peak shape suggests a volatile compound of mid-to-low polarity with a boiling point between of 130 – 150 °C (Figure45).

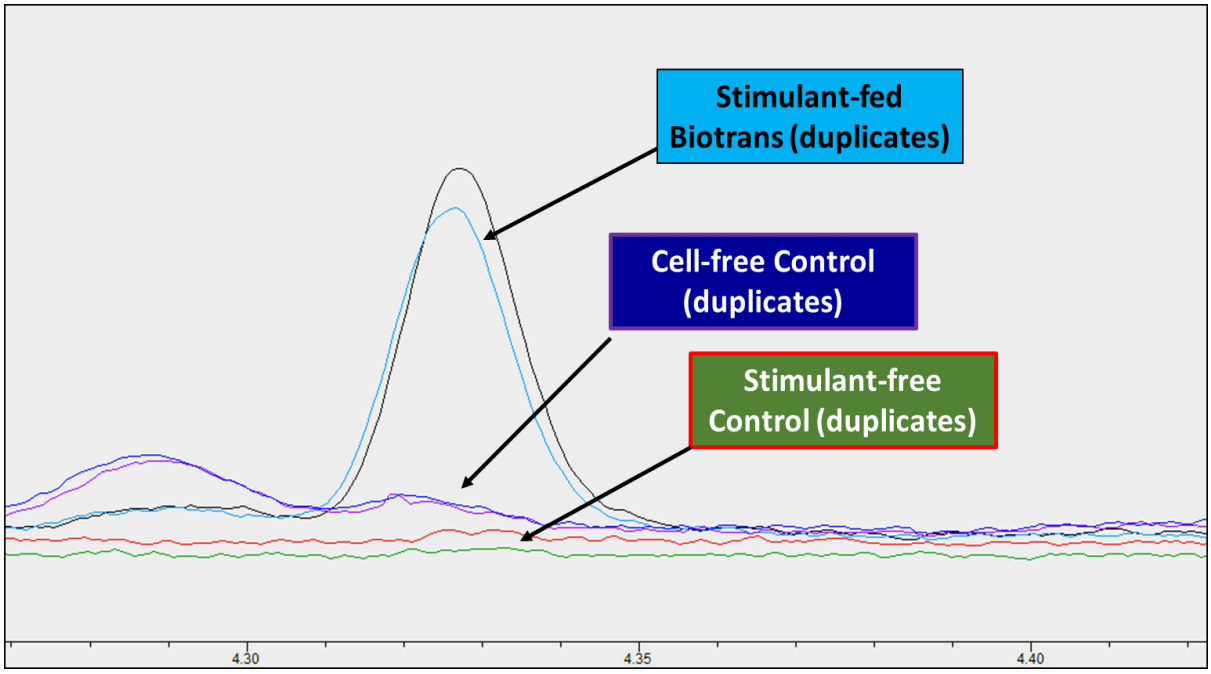

Figure 45:GC-FID trace of an unidentified analyte (retention time: 4.33 minutes) present in stimulant-fed *E. coli* at *t* = 21 hours.

A GC-FID method for the analysis of stimulants has not been developed, as many of the stimulants are involatile and unsuitable for GC analysis. However, as determined in the cell-free control experiments, five peaks correlating to different compounds in the stimulant cocktail are observable by GC in the stimulant-fed biotransformation; the metabolism of three of these stimulants over the course of the assay has also been documented as shown in Figure 46.

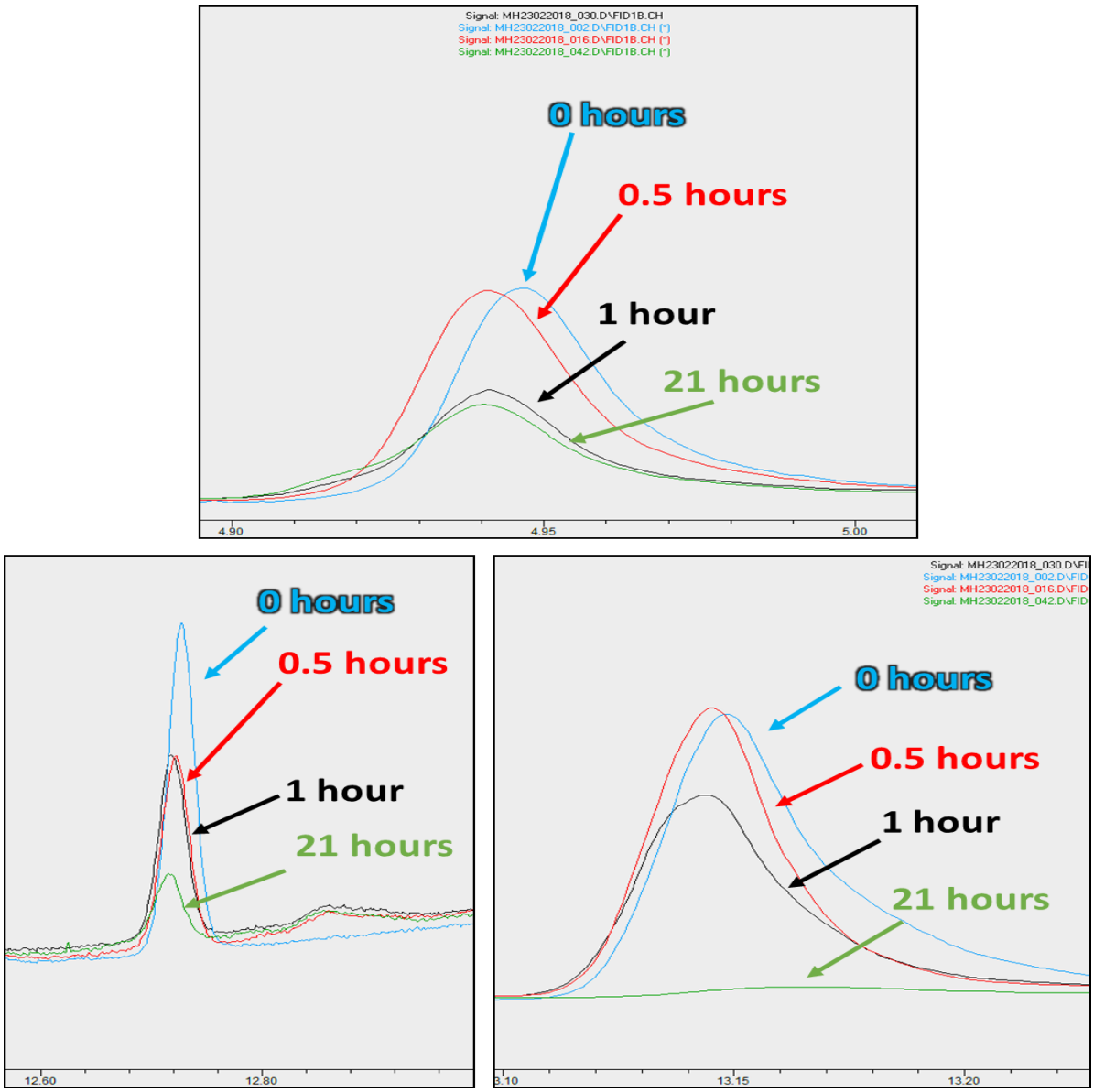

Figure 46:Stimulant consumption during *E. coli* stimulant-fed biotransformation. Peak retention times: 4.92, 12.72, 13.15 minutes.

8.2.4.2 *Lactobacillus Fermentum (NCIMB 11840)*

As with the *E. coli* experiment, the *L. fermentum* biomass was assayed in PBS solution containing the stimulant mix after a period of starvation. Relative to the *E. coli* experiment, no unique or identifiable peaks were observed under these conditions using *L. fermentum*,

and no significant differences in the chromatographic peak profile was observed when compared to the stimulant-free and cell-free controls. In addition, the *L. fermentum* strain did not appear to metabolise the same GC-detectable stimulants found in the *E. coli* experiment and cell-free controls, perhaps explaining the lack of biomarkers produced (Figure47).

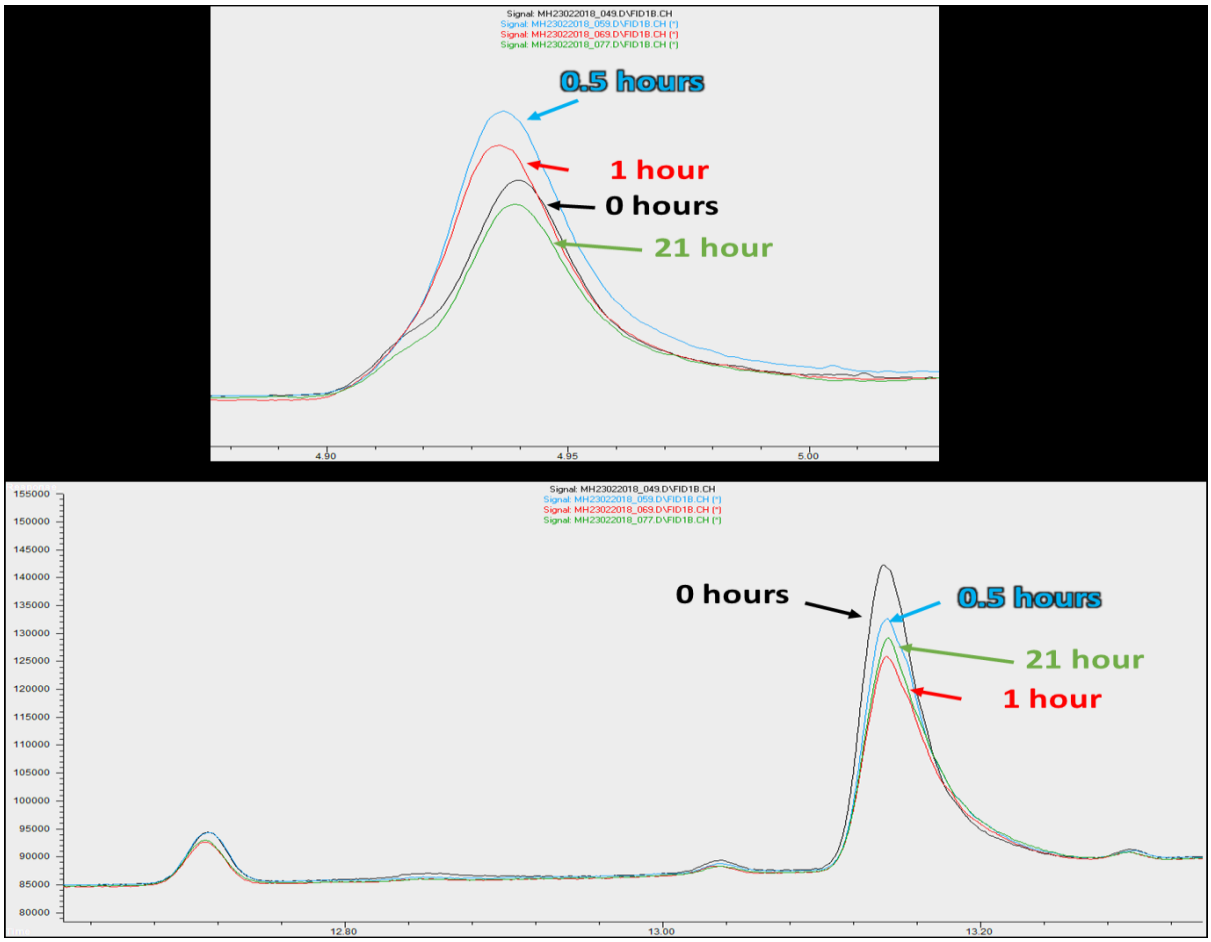

Figure 47:Lack of stimulant consumption during L. fermentum stimulant-fed biotransformation. Peak retention times: 4.92, 12.72, 13.15 minutes.

8.2.5 Discussion

In this pilot study, the preliminary data indicated that under micro-aerobic conditions, *E. coli* appears to metabolise components of the stimulant OSD to produce GC-FID-detectable organic compounds, which are absent in the control experiments under the same conditions. The data indicates a very close retention time match with an acetic acid reference standard. Two further unknown analyte peaks were also observed in the stimulant-fed biotransformations. It can be speculated, based on comparative retention times and boiling points to known standards, that one of the unknown analytes could be a low-polarity compound with a boiling point between 190 - 200 °C. The second unknown analyte appears after 21 hours into the biotransformation, and could be elucidated as a mid-low polarity compound with a boiling point between 130 - 150 °C. In these *E. coli* experiments, none of the remaining targeted phenols and fatty acids biomarkers were detected under these assay conditions. Furthermore, *L. fermentum* does not appear to produce any of the targeted biomarkers of interest.

This pilot study had several limitations. Compared to other analytical platforms, GC-FID lacks sufficient sensitivity to discern target VOCs and to characterise all peaks. In subsequent experiments VOC analysis was performed using GC-MS. Also to augment the VOC response, further biotransformations containing *E. coli* and *L. fermentum* were planned to be performed using increased biocatalyst loadings in the main high throughput study. Additional bacterial strains were also further investigated in the high throughput study.

Key outcomes for optimization in the high throughput and clinical study:

- More sensitive analytical technique to identify specific compounds
- The use of internal standards to confirm peak identity
- OSD optimisation to improve VOC response for culturing experiments and clinical assessment of patients.

8.2.6 Conclusion

Main findings of this preliminary experiment supported the hypothesis that under micro-aerobic conditions bacteria fed with selected substrates were capable of increased VOC production. Although no significant peaks were produced apart from acetic acid from Stimulant-fed *E. coli,* these preliminary experiments helped to establish the method for further culturing work that was intended to further investigate augmentation of VOC response.

8.3 Optimization of bacterial culture conditions and high throughput method of sample analysis by GC-MS

The purpose of this study was to further optimise the constituents of the stimulus mix and to verify its effects in cancer specific bacterial strains (Escherichia coli, Lactobacillus fermentum, Streptococcus salivarius, Klebsiella pneumoniae and Streptococcus anginosus). Dose response characteristics were also assessed. These experiments were intended to inform initial clinical trial in humans.

8.3.1 Aims

i. Optimising the composition of the stimulant mix to maximise VOC response.

ii. To augment VOC production by examining the metabolic pathways of upregulated bacteria present in oesophago-gastric cancer.

8.3.2 Methods

8.3.2.1 Stimulus mix production

Based on the results of the pilot study further refinement of the stimulant mix was undertaken in order to increase concentrations with the aim of improving the likelihood of observing elevated VOC responses from several bacterial strains. This section explains the constituents of the stimulant mix such that it is suitable for both culturing experiments and clinical assessment.

In comparison with the previous stimulus mix in the pilot study, some of the components were omitted from the recipe as food grade equivalents were either not available or had

long lead times for delivery including ethanol, xylose and lactic acid. Two additional components were added to the stimulus mix, citric acid and acetic acid, as these are common metabolites linked to central metabolism that could also be processed by the gut microbiome to augment a VOC response.

Concentrations of components used in the formulation were based on: i) Recommended Daily Intake (RDI) of a 70 Kg adult from a number of sources listed in table 25 below or ii) mix solubility limit. This enabled the same stimulus mix used for culturing work to be also given as an oral stimulant drink (OSD) for patients in the next section.

Component	Pilot study mix	Stimulus mix	RDI for healthy adult (70kg)	OSD (Stimulus mix equivalent)
	g/L	g/L	g/kg/day	g/kg
Tyrosine	0.192	1.75	0.025[1]	0.0025
Phenylalanine	0.596	-	-	
Glutamic Acid	0.469	21	0.035[2]	0.035
Glucose	0.576	130	1.857[3]	0.186
Xylose	0.458	-	-	-
Lactose	1.027	130	1.857[3]	0.186
Ethanol	0.136	-	-	-
Glycerol	0.280	193[4]	-	0.276
Sorbitol	0.552	35	1.857[3]	0.050
Citric acid	-	140	-	0.020
Acetic acid	-	70	-	0.010

Table 25**:**Stimulus mix and OSD composition summary, total bodyweight defined for 70 kg.

[1] FAO/WHO/UNU, *WHO Press.,* 2007, WHO Technical Report Series 935, 150

[2] European Food Standard Authority, *EFSA Journal*, 2017; **15**(7), 1-90

[3] Institute of Medicine, *Dietary Reference Intakes: The Essential Guide to Nutrient Requirements*, The National Academies Press. Washington, DC, 2006

[4] This value represents 29% lower glycerol content compared to Covonia cough syrup, http://www.covonia.co.uk/products/dry-tickly-cough-linctus.html (accessed March 2018)

8.3.2.2 *Bacterial Culture*

While each component of the OSD was formulated to not exceed its dietary RDI, a parallel high-throughput microbial growth/stimulus study was conducted *in vitro* to inform further development of the OSD towards an optimal composition for maximum clinical efficacy. Several study parameters were optimised during the work as described below, to enhance cell growth, culture sampling and VOC analysis.

Bacterial species to be cultured and analysed for VOC production during the *in vitro* component of this study were identified from the previous chapter (table26). Suitable strains were then obtained from culture collections such as NCIMB. All culture work was carried out in Cat1 or Cat2 laboratories at Ingenza as appropriate, according to UK microbiological regulatory guidelines using well-established cell culture protocols.

Organism	**NCIMB NUMBER**
Escherichia_Coli	NCIMB9552
Lactobacillus Fermentum	NCIMB11840
Streptococcus Salivarius	NCIMB701779
Klebsiella Pneumoniae	NCIMB13281
Streptococcus Anginosus	NCIMB702496

Table 26:Showing the bacterial strains and their commercial sourcing numbers (NCIMB).

Master and working cell banks of bacterial strains commonly associated with oesophago-gastric cancer were sourced from commercial cell depositories (NCIMB or ATCC), with all

associated testing for purity/viability of strain stocks. Strains were cultured overnight from working cell bank inocula, in liquid phase using the culture media for each strain recommended by the relevant cell depository, under aerobic conditions.

Following this, the cultures were back-diluted to OD_{600}=0.1 and grown aerobically until OD_{600}=0.6 was achieved. This ensured cells were in a metabolically active and viable state (exponential phase). Media was removed by centrifugation and washed with PBS (phosphate buffered saline). Cells were re-suspended in PBS and exposed to a 'starvation' phase under micro-aerobic conditions (nitrogen sparge). PBS did not contain any nutrients, so not to facilitate any growth. This was done to simulate the fasting process patients will undertake during clinical trials, and to convert the culture to an anaerobic state, to represent more closely the gastric environment.

PBS was removed by centrifugation and cells re-suspended in PBS (negative control), or PBS + stimulant mix as above. Flasks were sparged to maintain micro-aerobic conditions. Cell-free control was also produced (PBS + stimulant cocktail). Cells were incubated at 37 °C and samples taken at 1hr, 2hr and 16hr intervals for practicality. Timepoint samples were centrifuged and supernatant retained in 2 x 250 µL aliquots. For liquid samples, one aliquot was filtered (0.2 µM) and analysed by GC-FID using the previous method stated in the pilot study section.

Preliminary results proved that while *Escherichia coli* culture grew satisfactorily and generated detectable VOCs, all other cultures either did not achieve satisfactory growth under the initial protocol or did not produce detectable VOCs at the concentrations of stimuli used. It was therefore decided to revise the culture protocol to use growth media

and conditions that allowed greater culture biomass as well as to evaluate increased concentrations of potential VOC stimuli.

8.3.2.3 *Revised VOC stimulus to increase biomass of all strains*

Various culture media were tested to determine a suitable medium that significantly increased the growth rate and final biomass of all each strains, particularly *Streptococcus anginosus, Streptococcus salivarius, Klebsiella pneumoniae and Lactobacillus fermentum,* permitting sampling for VOCs to be conducted more effectively for each strain. First the strains were grown using commercially produced pre-made FAA plates subsequently using their recommended culture media to compare the growth rate.

Concentrations of stimulant mix components tyrosine, glutamic acid and glycerol were also significantly increased to 200 mM. The amino acid phenylalanine at 200 mM was also included in the stimulant mix as an additional precursor of the VOC phenol. The stimuli mix and composition of the stimulus assay, controls and spike standards are shown below in table 27. The VOC stimulus assay protocol is shown in table 28.

Condition	Stimuli present	PBS(µl)	Bacterial Culture
VOC stimulation of bacterial culture	L-tyr, L-phe, L-glu Glucose, Glycerol Sorbitol, Lactose, Xylose, Ethanol	146.23	500
Stimuli Free Control	No	1500	500
CFC Cell Free Control	L-tyr, L-phe, L-glu Glucose, Glycerol Sorbitol, Lactose, Xylose, Ethanol	646.23	N/A
Spike-standrd A 3-ethyl phenol	No	1480	500
Spike-standard B Hexanoic acid	No	1479	500

Table 27:Laboratory culture stimuli mix and stimulus assay composition.

Assay Conditions	
Growth medium	Glucose or Glycerol
Pre-culture	50 mL medium in 250 mL shake flask + 20 µl cell bank (200 µl for slow growers) 37°C, 250 rpm (up to 3 days for slow growers)
Growth	Back dilution to OD_{600} =0.1, 37°C, 250 rpm Final OD_{600} = 0.5-1.0
Starvation/biotransformation medium	Sterile PBS pH 5.0
Starvation conditions	20 mL resuspension N_2 sparging, 25 sec 1h, 37 °C, 250 rpm (anaerobic)
VOC stimulus assay	4.5-6.5 mL resuspension, N_2 sparging, 15 sec 37 °C, 250 rpm (anaerobic)
Analysis	Headspace sample to TD tube using pump
Time points	2 hour

Table 28:Laboratory culture VOC stimulus protocol.

8.3.2.4 *Revised culturing and analytical protocols to address limitations*

The original project plan anticipated the use of an OmniLog culture device to maximise throughput of culture conditions and VOC sampling. However, a number of difficulties were found with the use of the OmniLog. The slow growth rate of many of the cancer associated bacteria required vigorous agitation in flasks to generate biomass for VOC sampling, rendering the reduced agitation possible in the OmniLog insufficient. Also the critical need to maintain efficient vessel sealing during headspace sampling, to avoid loss or cross-contamination of highly volatile gases, could not be satisfactorily achieved with OmniLog plates, despite various seals being tested. Therefore a protocol for biomass generation was established in shake flask cultures followed by VOC stimulation in 50ml falcon tubes with final nitrogen sparging and VOC sampling in headspace vials, with the required accuracy and reproducibility. The protocol is shown in flowchart below (Figure48).

Following biomass generation, a culture starvation phase was conducted at pH 5.0 under micro-aerobic conditions prior to addition of the VOC stimulus mix (0.1 M concentration: tyrosine, glutamic acid, glucose, lactose, sorbitol, glycerol, ethanol, xylose, phenylalanine). These steps would help to establish a metabolic state in each organism that would facilitate the synthesis of the maximum number of biosynthetic enzymes and which would be closer to that of the organisms in the gastric environment. In this way the detection of VOCs under laboratory test conditions was considered to be a more accurate indicator of a similar response of the oesophago-gastric microbiome when exposed to stimulus mix. To further assess the impact of culture medium carbon source, a number of cultures were first assessed for VOC stimulation following growth on either glucose (catabolite repressing) or glycerol (non-catabolite repressing) as the carbon source.

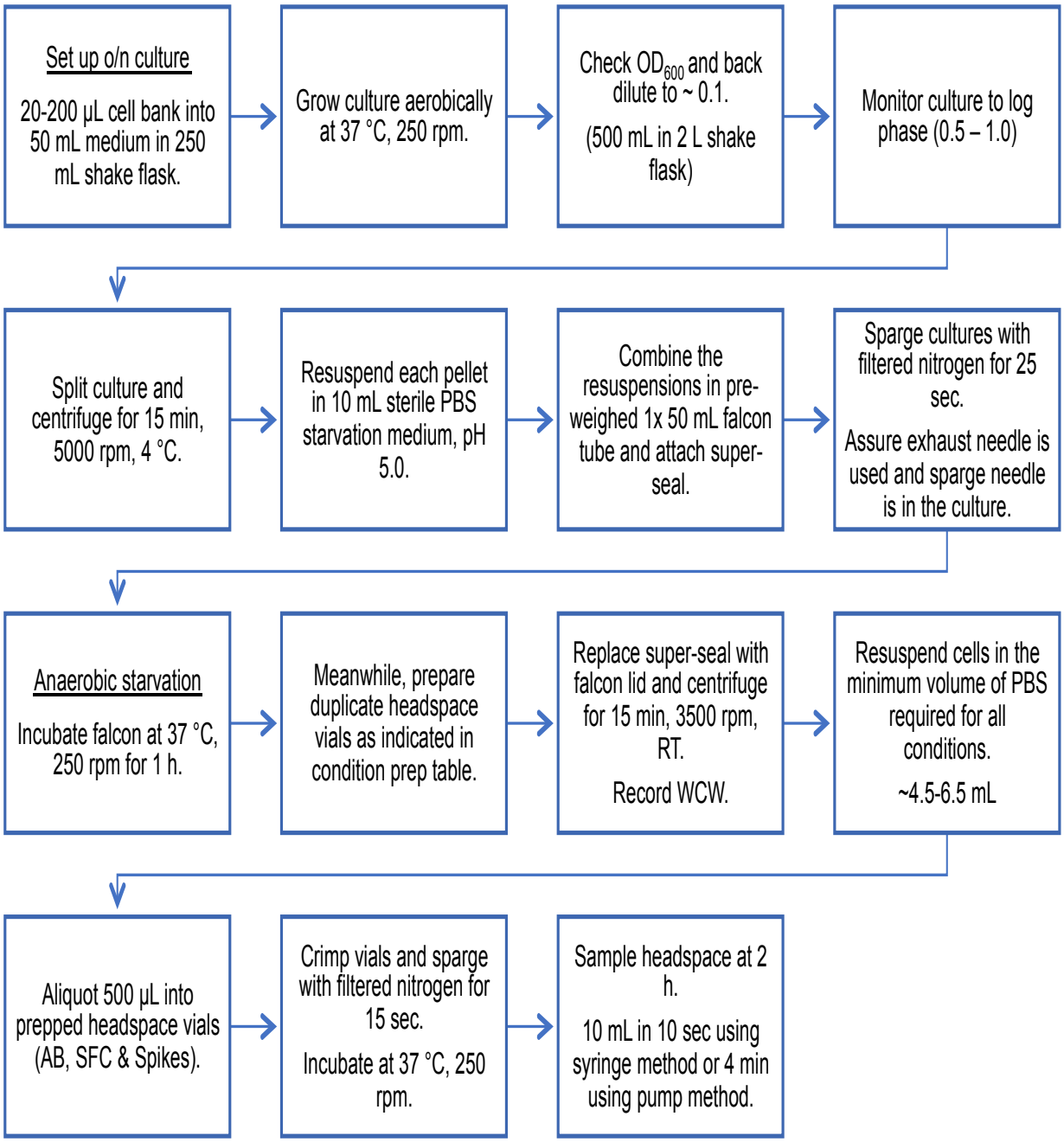

Figure 48:High-throughput in vitro culture stimulation and VOC sampling protocol.

8.3.2.5 *Mass spectrometry analysis*

VOC analysis by GC-FID in the pilot study was found to be significantly less sensitive than other analytical capabilities. TD-GC-MS was employed to qualify, with high accuracy, the identity of the VOCs within the samples.

The protocol for analysis of VOCs described herein was optimised for targeted analysis of specific compounds based on previous VOC studies in oesophago-gastric cancer. These include acetone, acetic acid, butyric acid, pentanoic acid, hexanoic acid, phenol and Acetaldehyde. Headspace VOCs were collected onto conditioned TD tubes for analysis. To permit additional throughput of experimental conditions, sample replicates and controls, it was also decided to limit samples to a 2hr timepoint.

A comprehensive theory of GC-MS is reported in the literature [84], however a brief explanation of the protocol used in this experiment is as follows. Inner-coated stainless steel Tenax/Carbograph-5TD sorbent tubes (pre-packed with 200 mg of Tenax and 100 mg of Carbograph5) were purchased from Markes International (Llantrisant, UK). Prior to sample collection, TD tubes were pre-conditioned using a Markes TC-20 tube conditioner at 325 °C for 40 min under nitrogen stream at 20 psi head pressure. Details of the conditions of analysis using TD-GC-MS have been published elsewhere [84]. In brief, sample of TD tube were pre-purged for 3 min under constant 50 mL/min helium prior to 280 °C desorption for 10 min. This followed by desorption of the cold trap (material emission U-T12ME-2S) under heating from 10 °C to 290 °C at 99 °C/min and held for 4 min. The flow path onto GC was heated constantly at 200 °C.

Using C tubing, 10mL of the headspace from culture samples was pumped for four minutes across thermal desorption tubes using a precision handheld pump (SKC Ltd, UK). Upon collection, samples were incubated at 37°C and analysed within four hours in accordance with the VOC stability.

The tubes were then analysed using a Agilent 7890B GC with 5977A MSD (Agilent Technologies,Cheshire, UK), coupling to a Markes TD-100 thermal desorption unit. VOC separation was explained in section 2.3.2.4.

8.3.2.6 *Data processing and statistical analysis*

Data acquired from Agilent GC-MS system was processed using MassHunter software version B.08 (Agilent Technologies) while compound identification was performed with matching of NIST Mass Spectral Library version 2.2 and retention indices of authentic standards. GraphPad Prism version 7 (La Jolla, CA) was used for further statistical analysis and graphical presentation.

8.3.3 Results

The overall assessment of the headspace data of 5 bacterial strains indicates there are some positive leads from the *in-vitro* work. Active biotransformation (AB) compared to controls comprising bacterial cultures that were stimulant-free (SFC) and stimulant compositions that were free of bacterial cultures (CFC). Tow internal standards were used for quality measures, Spike A and B containing 3-ethyl phenol and hexanoic acid. Data analysed by GC-MS is summarised below in table 29 showing the significant compounds produced by the microbes in comparison to SFC and CFC with specific conditions tested (Figure 49).

	Glycerol media	**Glucose media**
E. coli (NCIMB 9552)	Acetate, propanoic acid, butanoic acid	Hexanal, butanoic acid, pentanoic acid, hexanoic acid
L. fermentum (NCIMB 11840)	Acetaldehyde, ethyl phenol	Acetone
S. salivarius (NCIMB 701779)	Acetate, pentanoic acid,	Insufficient growth
S. anginosus (NCIMB 702496)	Acetone, Butanoic acid, Ethyl phenol	Butanoic acid
K. pneumonia (NCIMB 13281)	Hexanoic acid	Butanoic acid, pentanoic acid

Table 29:Elevated VOC levels in active biotransformation versus controls (Stimuli cocktail composition, all at 0.1 M concentration.

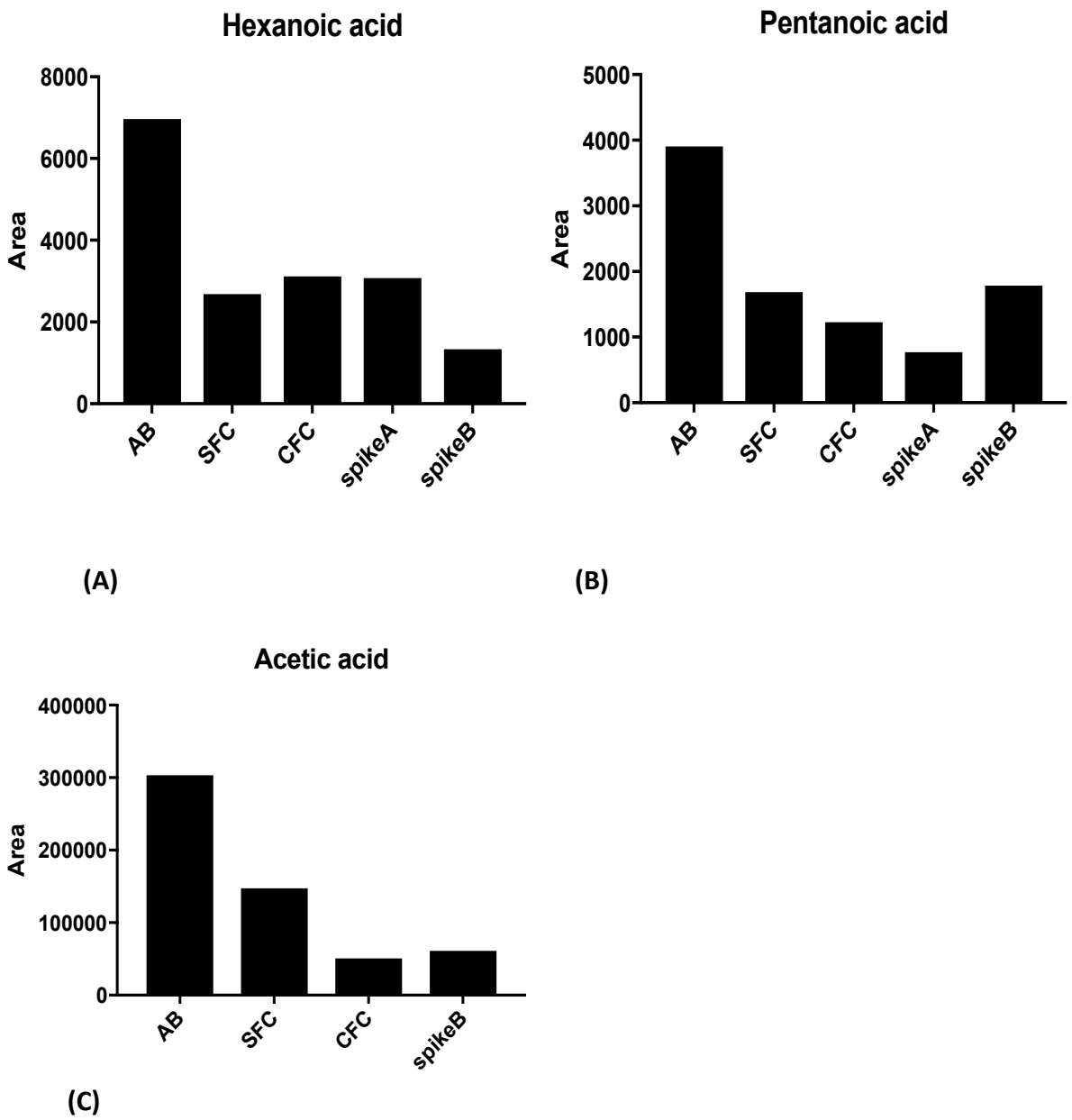

Figure 49:Examples of elevated VOCs in the headspace of **(A)** *Escherichia Coli in glucose media*, **(B)** *Klebsiella pneumonia in glucose media*, **(C)** *Streptococcus salivarius in glycerol media*. Active biotransformation (AB) compared to controls comprising bacterial cultures that were stimulant-free (SFC) and stimulant compositions that were free of bacterial cultures (CFC). Spike A and B contained 3-ethyl phenol and hexanoic acid as internal compound standards. Data is derived from VOC analysis by GC-MS.

8.3.4 Discussion

This specific high throughput study has maximised VOC response by optimising the composition of OSDs. The findings from in-vitro bacterial culture experiments provided evidence that the cancer associated bacteria is, at least in part, responsible for the observed changes in target VOC. The inherent mechanisms of genetic and biochemical regulation of the oesophago-gastric microbes under evaluation was considered important in the composition of growth media and the choice of carbon source used in the laboratory to generate cell biomass. Glucose for example exerts catabolite repression upon the cell, which inhibits enzymes necessary to catabolise carbon sources other than itself. Media that is rich in supplements such as amino acids or its metabolic intermediates also represses the biosynthesis of corresponding biosynthetic enzymes for these compounds, that are non-essential under these conditions but whose activity may be required for VOC production. Therefore, with the exception of the initial culture growth tests using supplier recommended media, the culture medium used throughout this work was a defined minimal salts medium lacking non-essential supplementation.

The challenge was to elicit short chain fatty acid response to stimulus mix from the oesophago-gastric cancer-associated microbiome. The biotransformations that were used were at pH 5 in phosphate buffered saline. This pH is around the pKa values for short chain fatty acids e.g. acetic acid pKa = 4.8, propanoic acid = 4.87, butanoic acid = 4.82. This means approximately 50% of the acid will be in the acid form and 50% in the salt form. The salt form is much less volatile so in theory there could be an underestimation of the total amount of short chain fatty acids in each sample by only measuring the VOC's in the

headspace. However, they have demonstrated 'proof of principal' that the exhaled biomarkers of oesophago-gastric cancer may be augmented by an OSD.

The culturing and stimulus work was essential to define the optimal performance of the OSD in augmenting VOC response to most effectively correlate with the response of patients with oesophago-gastric cancer. Initially, the composition and concentration of individual OSD constituents for *in vitro* study closely replicated that of the OSD prepared for clinical use. However, in further *in vitro* analyses, constituent concentrations were then significantly increased, permitting a much broader assessment of individual stimulant thresholds, temporal profiles and concerted effects of stimuli upon the extent of microbial VOC production. These outputs can still inform the design of subsequent patient dosage studies, recognising that many foods and common nutritional supplements (e.g. vitamins, minerals, amino acids) often greatly exceed the RDAs for compounds potentially suitable in the OSD.

The composition and degree of VOC generation from these OSD stimulated cultures versus non-stimulated controls has informed the choice of safe-dosage OSD formulations, to be used in the development of clinical studies discussed in the next section.

8.3.5 Conclusion

Previous work has shown that factors affecting the gut microenvironment, including the microbiome, can significantly affect volatile organic compounds within exhaled breath. This high throughput study enhanced the opportunity to intentionally modify the gut microbial environment in a targeted way to derive mechanistic insight in to VOC production relevant to early oesophago-gastric cancer detection.

8.4 Clinical study

8.4.1

The intention of this interventional case-control study was to profile VOCs in response to Oral Stimulant Drink (OSD) given to cancer patients compared to control subjects. Dominant VOCs that differentiate cancer from non-cancer in response to OSDs will provide the basis for targeted in-vitro experiments.

8.4.2 Aims

i. To develop a process for formulating OSD for human consumption.

ii. To assess VOC response to OSD administration in oesophago-gastric cancer patients and healthy control subjects.

8.4.3 Methods

8.4.3.1 *Oral stimulant drink formulation process*

Process Risk Assessment determined that the stimulant mix manufacture be removed from the laboratory environment. Time in a commercial kitchen local to Ingenza was hired for an Oral Stimulant Drink (OSD) manufacture. The OSD was then prepared to principles of good manufacturing practice and quality Management System with full traceability and batch records. A final certificate was produced with completion of each batch. All OSD components and consumables were sourced as either “food” or “pharmaceutical” grade.

This was to ensure no contamination occurred and that the drink was also fit for human consumption. Drinking water was used as the diluent in the OSD, procured from a commercial source.

A 100 mL solution was prepared per patient, either with OSD (concentrations described in section 1.3.2.1) or water blank for the control group. 0.2 µM filtration and final fill was performed aseptically to ensure minimal contamination from environmental microbes. Ethical, safety and acceptability considerations of OSD components were performed, such as normal dietary presence, recommended daily allowance (RDA) and palatability. Storage and shipment were under controlled refrigerated temperatures (2-8°C). Accelerated stability studies were performed to provide an expiry date of the OSD batches.

8.4.3.2 *Patient selection and exhaled breath sample collection*

Patients with oesophago-gastric cancer as well as healthy subjects with no history of upper gastrointestinal disease were recruited at the time of routine outpatient assessment. Patients were required to fast for a minimum of 6 hours prior to breath sampling. Patient's exhaled breath samples were collected using Nalophan (Kalle UK Ltd, Witham, UK) bags produced in a standardised fashion (17). Each bag was formed from double thickness (2x25µm) nalophan sheets (31x21cm) with the ends sealed with a hand operated impulse heat sealer (Mercier Corporation, New Taipei City, Taiwan) twice over omitting a single aperture for the insertion of a 1mL Luer-Lok tipped syringe (Becton Dickinson, New Jersey, USA). Cable ties were used to secure the syringe to ensure airtightness was achieved whilst still allowing for a means to inflate the bag. Rigorous quality control was conducted ensuring

each bag had a fixed volume of 1.5L and was washed out with dry synthetic air (BOC, Guilford, UK) following assembly.

Breath samples were collected from patients who were seated and at rest. Patients were instructed to undertake a complete nasal inhalation followed by a complete oral exhalation of mixed alveolar breath into the bag via the syringe barrel. Exhalation was halted once the bag was at maximum capacity ensuring a small surface (film) to volume (sample) ratio to maximise the VOCs stability. The plunger from the syringe was then replaced, sealing the sample bag.

A baseline breath sample was collected at the start of the study period. Participants were then asked to consume an oral stimulant drink (OSD). Following consumption of the OSD participants were asked to rinse their mouth with water in order to eliminate any oral residue of the OSD. Further breath samples were then collected at 30 and 60 minutes following ingestion of the OSD.

Nalophan bag samples were then pumped into conditioned TD sorbent tubes utilising a hand held pocket pump (SKC Ltd, Dorset, UK) transferring breath samples at a constant rate of 50 mL/min for 10 minutes to collect a total of 500 mL of breath onto the tubes. This protocol also required a 21G Agani needle (Terumo, New Jersey, USA) and 5 mm plastic tubing to provide the means to puncture the bag and provide a route for the breath sample to be pumped onto the tube. Upon collection, samples were incubated at 37°C and analysed within four hours in accordance with the VOC stability within the bags.

Inclusion criteria: Patients with the following characteristics were eligible for inclusion in this study:

[1] Male and females aged 18-90 years

[2] Patients with biopsy proven cancer of the oesophagus or stomach

[3] Control subjects without a history of upper gastrointestinal tract disease

Exclusion criteria: Patients and controls with the following characteristics were not be eligible for inclusion in this study:

[1] Individuals who are unable or unwilling to provide informed written consent

[2] Individuals who have received antibiotic therapy within the last 4 weeks

[3] Individuals who have undergone previous gastrointestinal resection(s)

[4] Individuals with allergies to any of the constituents of the OSD drink

Local ethics committee approval was granted for this study (Ref: 18/LO/0078) and written informed consent was obtained from all patients prior to enrolment in the study.

8.4.3.3 *VOC analysis of mass spectrometry*

The protocol for analysis of VOCs described herein was similar to the compounds analysed in the previous culturing section 5.3.2.5. Analysis was conducted using two mass spectrometry techniques. Proton Transfer Reaction Time of Flight Mass Spectrometry (PTR-ToF-MS 100 ionicon Analytik, Innsbruck, Austria) using its direct injection mass spectrometric technique was employed for accurately quantifying VOCs in exhaled breath samples to the degree of parts-per-trillion per volume (*pptv*). Selective reagent ionization

was employed to cycle between H3O+ and NO+ reagent ions to maximise the ionizing capacity of multiple VOCs.

The PTR-ToF-MS instrument was coupled with a TD customised auto sampler unit (TD100-xr, Markes Ltd, LIantrisant UK). The comprehensive theory surrounding PTR-ToF-MS has been described elsewhere [64], however a basic overview and the protocol used for this study are as follows. Analysis was performed with a one-stage desorption method; tubes were desorbed for 10 minutes at 280°C using 1300 sccm of nitrogen, purified using a Supelpure HC hydrocarbon trap (Sigma-Aldrich, 131 St. Louis, MO). VOCs were then transferred to the PTR unit through an inlet made by peek tubing kept at 110°C , with a flow rate of 130 sccm. Drift tube analysis conditions were the following: temperature 110°C pressure 2.30 mbar, and voltage 350 V, resulting in an E/N of 84 Td. Quantification of VOCs was averaged over a 30 and 40 seconds period for each ion respectively with the time discrepancy adjusting for ion switching. Produced ions were separated and detected based on ion migration determined by mass-to-charge ratio (time of flight).

Replicates of the same breath samples placed in the thermal desorption tubes were also analysed by GC-MS, method explained in previous section 2.3.2.4. Cross-platform validation between the GC-MS and the PTR-MS data was conducted to accurately identify the quantified compounds measured via PTR-MS.

8.4.3.4 *Data processing and statistical analysis*

PTR-ToF-MS data were extracted using PTR-MS viewer version 3.2.8.0 (Ionicon Analytik) and analysed using pre-processed scripts written with R programming language. GC-MS data were extracted using MassHunter software version B.07 SP1 (Agilent Technologies) and further analysis was conducted using a custom designed in-house built software MassHub. VOCs peaks identification was performed using NIST mass spectral library (National Institute of Standards and Technology version 2.0). VOC levels were compared by average fold change. Cancer patients were also compared based on chemotherapeutic status. The Mann Whitney U test was used to compare statistical differences using GraphPad Prism (version 7.04 for Windows, GraphPad Software, La Jolla California USA).

8.4.3.5 *Quality Control*

Quality control processes were performed for the duration of the project on a daily basis. For PTR-ToF-MS, two types of quality checks were performed every day to ensure a constant analytical performance[37]. A first quality control (instrument quality control) evaluated instrument reproducibility with three ionisation modes against four parameters: impurities, fragmentation, mass resolution and accuracy. The permeation unit was directly connected to the PTR-MS inlet through a tube peek union connection. Measurement was carried out for five minutes. Accuracy was evaluated through quantification of benzene certified standard permeation tube (Kin-Tek Analytical Inc., La Marque TX). The second quality control (standard quality control) method evaluated the recovery of VOCs from desorbed Thermal Desorption tubes (TDt). TDt were loaded with standard mix from the permeation unit with a previously described method [63]. Concentrations of VOCs calculated

from TDt were compared to concentrations calculated from the permeation unit through direct measurement obtained during the first quality control procedure. The first quality check evaluated the performance of the PTR-MS instrument, while the second assessed the performance of TD unit coupled with the PTR-MS.

As for GC-MS, a similar quality control procedure was adopted. Five TDt were loaded daily with VOCs standard mixture from the permeation unit and were analysed. Retention time and peak area was performed if the relative standard deviation was below 5%. To quantify VOCs, calibration curves were performed with the use of the permeation unit, modulating the flow load TDt at different concentrations.

8.4.4 Results

Thirty patients with oesophago-gastric cancer and 30 control subjects were recruited. Patients with gastric and oesophageal cancer were older (MWU, p=0.017) and predominantly male (p=0.029) compared to control patients. All participants were able to consume the OSD and there were no observed or reported adverse events. Exhaled acetone, a marker of ketosis (a state of energy depletion) was assessed in order to verify the administration of a nutritional stimulus, in the form of the OSD. Acetone levels in both cancer and control subjects decreased following ingestion of the OSD confirming nutritional stimulation that occurred.

Following ingestion of the OSD, target VOCs in cancer patients were detected at higher levels as indicated by the average fold change in VOC concentrations at 30 and 60 minutes (table 30,31). With the exception of butyric acid (30 minutes time point), control subjects exhibited a ≤10% variation in target VOC levels following ingestion of the OSD. Mean fold change variation in exhaled hexanoic acid, and pentanoic acid. There was broad agreement between the PTR-MS and GC-MS data.

For cancer patients who had previously received chemoradiotherapy, OSD response for pentatonic appeared to be suppressed such that it was similar to control subjects (Figure 50).

	Cancer			Controls		
	0 mins	**30 mins**	**60 mins**	**0 mins**	**30 mins**	**60 mins**
Acetone	1.0	1.0	0.9	1.0	1.0	0.9
Acetic Acid	1.0	1.1	1.2	1.0	1.0	0.9
Butyric Acid	1.0	1.6	1.3	1.0	1.2	1.0
Pentanoic Acid	1.0	1.1	1.2	1.0	0.9	1.0
Hexanoic Acid	1.0	1.1	1.2	1.0	1.0	1.0
Phenol	1.0	1.2	1.3	1.0	1.0	0.9
Ethyl Phenol	1.0	1.2	1.3	1.0	0.9	1.0
Acetaldehyde	1.0	1.1	1.2	1.0	1.1	1.1

Table 30:Mean fold change in select exhaled VOCs following administration of oral stimulant drink. Data is derived from breath samples analysed by PTR-TOF-MS.

	Cancer			Controls		
	0 mins	**30 mins**	**60 mins**	**0 mins**	**30 mins**	**60 mins**
Acetone	1.0	1.1	1.0	1.0	1.0	0.9
Acetic Acid	1.0	1.1	1.2	1.0	1.0	0.9
Butyric Acid	1.0	1.6	1.3	1.0	1.1	0.9
Pentanoic Acid	1.0	1.0	1.3	1.0	1.1	1.0
Hexanoic Acid	1.0	1.1	1.2	1.0	1.0	1.0
Phenol	1.0	1.2	1.4	1.0	1.0	0.9
Ethyl Phenol	1.0	1.2	1.3	1.0	0.9	1.0
Acetaldehyde	1.0	1.1	1.2	1.0	1.1	1.1

Table 31:Mean fold change in select exhaled VOCs following administration of oral stimulant drink. Data is derived from breath samples analysed by GC-MS.

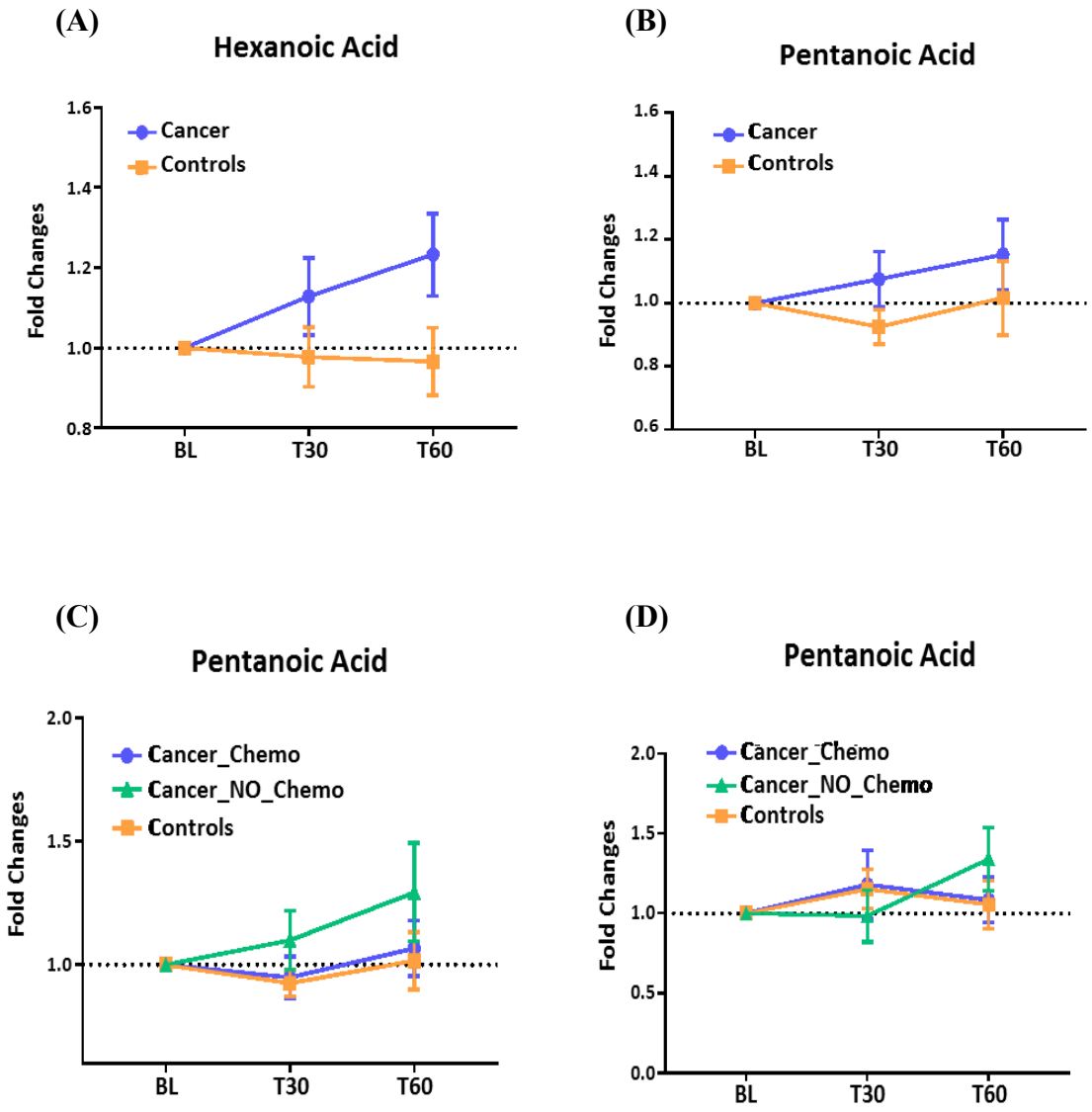

Figure 50:Mean fold change variation in exhaled (**A**) hexanoic acid and (**B-D**) pentanoic acid in both oesophago-gastric cancer patients and control subjects following ingestion of an oral stimulant drink. For pentanoic acid, cancer patients have been further sub classified according to chemoradiotherapy exposure (**C and D**). Breath samples were analysed by both PTR-ToF-MS (**A-C**) and GC-MS (**D**).

8.4.5 Discussion

Previous studies have provided verification that the tumour and proximate environment is responsible for unregulated production of target VOCs. The proposed innovation was made to improve the accuracy of a cancer breath test by testing whether nutritional substrates found within a 'normal' diet can augment the production of cancer associated VOCs.

The study demonstrated an increase in the generation of VOCs in patients with oesophago-gastric (OG) cancer in comparison to non-cancer subjects in response to an oral stimulant drink. Similarly, we have demonstrated in the previous section that bacteria commonly associated with oesophago-gastric cancer generate the same VOCs in laboratory testing, in response to many of the same metabolic stimuli. The following data formulated a microbial culture medium supplement that, under laboratory conditions, consistently elevated the production of VOCs, particularly pentanoic and hexanoic acids, from bacterial strains known to prominently associate with gastric tumours. Production of these same compounds was also significantly elevated in breath samples of oesophago-gastric (OG) cancer patients, compared to healthy subjects who had all consumed these same supplements.

The key to improving cancer-survival is earlier diagnosis, however symptoms are non-specific and commonly shared with benign diseases. There is a substantial commercial opportunity for a non-invasive test to triage patients for endoscopy. Cancer burden and unnecessary investigations of patients with non-specific symptoms result in substantial costs. The results provided a strong feasibility data towards a robust laboratory model that enables rapid ex-vivo optimisation of OSD formulation to yield high statistical correlation of patient VOC response and cancer presence. Also, an OSD that consistently elicits clear and early patient VOC markers of oesophago-gastric cancer.

The findings have also indicated that this response may be suppressed by the prior chemoradiotherapy, which is known to modify the intestinal microbiome [114].

This study has a number of important limitations that principally relates to the sampling method. Whilst it was intended to correlate between the VOCs produced in vitro and patients in response to specific stimuli, different substrate concentrations were used in the OSD and in vitro stimulus mix, with the latter having higher concentrations. The OSD and in vitro stimulus mix both however contained the same substrates. Although the culturing work was performed under anaerobic and mildly acidic conditions, to best simulate the gastric environment, there still remains a need to optimise these conditions with an enhanced simulated environment. One possible approach could be the use of a bioreactor.

Further work is now required to establish the optimal conditions for administering an augmented cancer breath test and to define potentially confounding factors and to fully demonstrate dose-response of the OSD, reproducibility of samples and robustness under likely clinical application.

8.4.6 Conclusion

In conclusion, this preliminary work has proven promising data showing specific variations of volatile compounds between exhaled breath of cancer and control patients. This work will help to facilitate a regulatory compliant, transferable manufacturing protocol to enable scale-up for further trials, trialling and commercialisation of OSD production for a VOC based cancer triage test.

8.5 Stakeholder analysis

8.5.1 Introduction

The current study explores the benefits of AMBEC as a new diagnostic test device and investigating its potential impact on the patient's journey. The intended impact of this test is to aid decision making and interventions that are undertaken as a result earlier disease detection.

Previously researchers have suggested the potential of breath analysis in the non-invasive diagnosis of several disease states. The implementation of this augmented test into clinical practice is a complex interaction of validation and reliability of the breath test, usability of the breath test and its acceptability to practicing medical professionals and patients.

8.5.2 Aims

The objective of this event was to solicit the opinion of different stakeholders on the utility of the test and where it stands according to patient pathway.

8.5.3 Methods

A three-hour stakeholder analysis event was undertaken at St Mary's hospital. We invited a multidisciplinary panel including clinicians from both primary and tertiary care, NHS commissioners, scientists and lay members. Panel members all had prior experience of oesophago-gastric cancer and its diagnostic pathway. All panel members provided informed written consent, giving permission for their voices to be recorded (Philips DPM600 digital pocket memo), before participating in the workshop.

Following a brief introduction to the day, a presentation covering the background to oesophago-gastric cancer including the current diagnostic pathway, which was given by a specialist trainee in general surgery with an upper gastrointestinal sub-speciality interest. Data were presented about oesophago-gastric cancer, noting that the incidence of oesophageal adenocarcinoma in the UK is the highest in the world. It was explained that patients in the UK take too long to present to their GP, partly because symptoms of heartburn and indigestion are not widely recognised as potentially indicative of oesophago-gastric cancer. Endoscopy is the gold standard for diagnosis, but it is costly and invasive and experienced as a very unpleasant investigation by many patients. Participants were given the opportunity to ask questions after the presentation. The presentation was followed by the workshop breaking down into small groups (5-10 people) for facilitated table discussions regarding the current diagnostic pathway. The groups were then brought back together and table discussions were summarised. This introduced the concept of the augmented breath test with oral stimulant drink, current evidence and the group's future ambitions.

The workshop discussions were recorded for subsequent analysis. Individual interviews with attendees were conducted during the session to capture their workshop experiences. Workshop recordings were transcribed and emergent thematic analysis undertaken. Two analyses were undertaken: (i) an analysis of patient feedback regarding the augmented breath test device, (ii) diagnostic strategy and clinical pathway.

8.5.4 Results

On 6th of June 2018, 23 people gathered at St Mary's hospital to provide guidance for the development of an augmented microbiome breath test. In respect to the diagnosis of oesophago-gastric cancer and the proposed new augmented breath test and diagnostic test strategy, five themes emerged from the workshop: awareness of oesophago-gastric cancer, values in testing and diagnosis, challenges arising from new test device, usability issues of test device. The co-construction of knowledge and summarised output from the shared experiences of members from the event are provided in the table below.

Key Themes Identifies	Summary	Quotations from transcripts
Awareness of oesophago-gastric cancer	There was general agreement regarding the lack of awareness of oesophago-gastric cancer, it's symptoms and the poor outcome. There was a strong feeling of a need for awareness.	*"It needs to be more like colorectal cancer because people know if they have blood in their stools it's a worrying sign and straight away they go to their GP."*
Values to testing and diagnosis	There were many barriers highlighted to the current diagnosis of oesophago-gastric cancer; these included the anxiety, invasiveness, cost and complications of endoscopy, lack of education, cultural reluctance to seek medical advice particularly amongst men in at risk age groups, difficulty in accessing primary care services, delayed recognition of potential cancer and subsequent referral by GPs and hospital delays for investigation.	*"The added value will be obvious after having the oral stimulant drink and what they would gain from usability up to high sensitivity levels."* *"How invasive the current test is could put a lot of people off."*
Challenges arising from the new test device	The practitioner delivering the result should be appropriately trained to explain the meaning of the result and council regarding further management and privacy and support need to be provided. Written information was not felt to be sufficient and the inclusion of a nurse was highlighted as beneficial.	*"It is important that people are trained to give out the results appropriately"* *" In terms of the location of the test, GP practice managers preferred the test to be done at the laboratory to avoid any issues with calibration and equipment maintenance"*
Usability issues of novel test device	There was some concern that the drink should taste good. Also a positive breath test would increase anxiety preceding endoscopy but this was felt to be unavoidable and would always occur in cancer diagnostic pathways and may in fact increase the uptake of subsequent endoscopy. It was important that patients with a negative breath test know to return to their GP if symptoms do not improve.	*" The drink will elicit a taste response immediately but we may wish to deliver it through a simple capsule to overcome potential taste issues"* *"Something before the invasive endoscopy test would encourage people more"* *"To satisfy criteria for screening the test should be cheap, effective and accessible''*

Table 32:Summary output from the shared experiences of members from the event.

8.5.5 Discussion

The workshop confirmed clinician and patient acceptability of the proposed breath testing strategy for cancer diagnosis and reinforced the lack of awareness that exists in regards to oesophago-gastric cancer and its symptoms among the general public. Breath testing would be ideally situated in the primary care setting. General practitioners within primary care are most often the first point of contact for patients with upper gastrointestinal symptoms and as such are primary source of referral for diagnostic testing. General practitioners and allied healthcare professionals within primary care could administer the augmented breath test in order to support decision making and triage patients for endoscopy.

A further important finding from this workshop was the majority of participants were unsure as to the optimal patient interface for any breath test in the future. This is clearly needed to plan in the next stage of the breath test development, as a robust breath test in clinical practice is unlikely to gain widespread dissemination using breath bags for sample collection. Hand-held devices, which are sensor based, would be better suited to the primary care setting and could be easily operated by a trained medical professional that can deliver the results in a professional manner to avoid patient anxiety.

The data obtained from undertaking this stakeholder event also provided feedback on research activities to date and will help guide future undertakings. Indeed, the output from the workshop have already influenced the research group's future research activities for the augmented VOC breath test for oesophago-gastric cancer as follows: identifying a broader scope of where the test could be located (including general practitioners, pharmacies and

workplaces), how the test result is presented to manage associated patient anxiety and the migration from currently used breath collection bags to sensor technology.

9 FUTURE WORK

This thesis has undertaken a systematic investigation of the microbiome and its influence to volatile compound production. It has investigated how the metabolism of upregulated bacteria in oesophago-gastric cancer can be used to augment breath testing as a noninvasive diagnostic tool. Through this work and from our previous research, we have found evidence of upregulated volatile fatty acids and phenol production in patients with oesophago-gastric cancers. Whilst the precise origin of these compounds remains unknown, this thesis has postulated that the cancer associated microbiome (*E.coli*, L fermentum, S salivarius, *K. pneumoniae* and *S. anginosus* and fusobacteria) may play an important role in their production.

The work described in the thesis has shown that the exposure of standard cultures of oesophago-gastric cancer associated microbiome to a mixture of judiciously selected of metabolic substrates resulted in an increase in headspace levels of the volatile fatty acids: butanoic, pentanoic and hexanoic acid. Whilst these findings give support to the hypothesis that cancer-associated bacteria are at least partly responsible for elevated VFA production, it largely neglects the complex onco-microbial environment that occurs in vivo. Therefore, It will be necessary for future work to validate the metabolic effect of predominant bacteria and seek leverage of the in vivo onco-microbial axis in order to address the specific challenges of early detection of oesophago-gastric cancer through the detection of VFAs and phenol within exhaled breath.

Also, through an understanding of the mechanism and kinetic of upregulated compound production in oesophago-gastric cancer, it may be possible to further augment this signal to

aid diagnostic performance. In order to achieve this, there will be a need to provide mechanistic proof of principal for the association between bacterial species and VFA production.

Further studies are needed to explore the complex relationship that occurs in-vivo between the cancer microbiome and metabolic pathways of VOC production using advanced ex vivo models including bioreactors. A bioreactor is a device that supports a biologically active environment such as within a human stomach. Under optimal conditions, cells and micro-organisms are able to perform their desired metabolic function with minimal impurities. Furthermore, important environmental factors such as temperature, pH, nutrient concentrations and dissolved gas levels can be precisely controlled. For these reasons a bioreactor offers an opportunity to establish a highly reliable *ex vivo* model of the tumour environment in which to investigate and validate our existing understanding of VOC production in oesophago-gastric cancer. The bioreactor fermenter will be equipped with pO_2, pH and temperature sensors, variable and fixed speed pumps to control the flow of HCl, $NaHCO_3$, bile acids, feeding medium, and the culture volume. The temperature in the bioreactor will be kept at 37±0.1 °C, and anaerobic conditions will be maintained by purging with nitrogen. In order to simulate the survival conditions of specific cancer associated microbiome, 100 mL of 0.01 M HCl will be added into the vessel to imitate the empty stomach, with pH value near 2. At the beginning of the experiment, 200 mL of media containing 107–109 cfu/mL of specific bacteria will be pumped into the vessel. Off gas analysis will be performed by transferring headspace gases, via a hand held precision pump on to thermal desorption tubes to detect VOC changes with respect to confounders used.

Part of future work should also include addressing the technical challenges, which existed in establishing the optimal OSD formulation that elicited a positive VOC response. The high-throughput *ex-vivo* culture, supplementation and VOC sampling protocols established in AMBEC study will require further enhancements based on patient sampling. To avoid underestimation of the total amount of short chain fatty acids measured in the headspace of culturing samples, the pH would be adjusted to pH < 3 just before the samples taken (by adding HCl for example) leading to more than 99% to be in the acidic form. This would enable an increase of acid volatility and potentially make a significant difference to the responses seen by GC. Also as gas solubility is lower at elevated temperatures, further work would be needed to improve the partition from liquid to gas phase by heating samples.

There will be a need to increase patient sample sizes to more closely monitor reproducibility of the proposed augmented breath test. The current OSD established will initiate a patient development trial in which the composition, concentration and relative balance of OSD constituents will be iteratively modified based on patient responses. Specific OSD variations eliciting the most significant and consistent patient responses will guide the focus in design-of-experiment laboratory protocols to assess various corresponding media supplementation experiments in each cycle of iteration, thereby accelerating optimisation of the OSD. The future output would be a single OSD that performs most consistently and acceptably across the widest patient sample. Bacterial culture and sampling parameters relevant to the gastric environment will also be varied, including sampling time, pH, anaerobiosis and simulated fasting.

This work has also established the presence of phenol-producing bacteria within the stomach, knowing from previous work that tyrosine appears to be converted to phenol within gastric

juice. To definitively establish the presence and functional action of the tyrosine *phenol-lyase* gene within the stomach it will be necessary to develop an assay to measure this enzyme specifically. This could be performed by isolating bacterial DNA from tissue samples and using qPCR for this enzyme found in tissue samples or gastric juice. However, this technique is challenging as the genetic sequence of this enzyme is not well-preserved across bacterial phylum and different genera. This means a single qPCR primer cannot be built to incorporate all forms of this gene in different bacteria. Another potential approach may be to attempt to isolate or purify the tyrosine phenol-lyase protein, and determine the primary amino acid sequence of the enzyme to help identify the responsible microbe for production of this enzyme. It will be possible to apply these approaches to stomach tissue samples as well as gastric juice samples. This would establish whether the effect of tyrosine phenol-lyase is greatest within the stomach mucosal tissues or gastric juice itself.

In summary, the overall aim of this thesis was to establish the mechanism responsible for specific volatile compound production on oesophago-gastric patients and how this is linked to microbial metabolism. It was possible to identify various bacteria that were upregulated in cancer patients and from these, specific ones were targeted to the production of volatile fatty acids and phenol. The work in the thesis has also demonstrated that specific nutrients affecting the gut microenvironment, including the microbiome, can influence volatile organic compounds within exhaled breath. The opportunity to intentionally modify the gut microbial environment in a targeted way may derive mechanistic insight into volatile fatty acid production relevant to early oesophago-gastric cancer detection.

10 REFERENCES

(1) Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int. J. Cancer.2010; 127, 2893–2917.

(2) Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. Gut. 2015; 64(3): 381–7.

(3) Pohl H, Sirovich B, Welch HG. Esophageal adenocarcinoma incidence: are we reaching the peak? Cancer Epidemiol Biomarkers Prev. 2010; 19, 1468–1470.

(4) Lepage C, Rachet B, Jooste V, Faivre J, Coleman MP. Continuing rapid increase in esophageal adenocarcinoma in England and Wales. Am. J. Gastroenterol. 2008; 103, 2694–2699.

(5) Steevens J, Botterweck AA, Dirx MJ, van den Brandt PA, Schouten LJ. Trends in incidence of oesophageal and stomach cancer subtypes in Europe. Eur. J. Gastroenterol. Hepatol. 2010; 22, 669–678.

(6) Bosetti C, Levi F, Ferlay J, Garavello W, Lucchini F, Bertuccio P, Negri E, La Vecchia C. Trends in oesophageal cancer incidence and mortality in Europe. Int. J. Cancer. 2008; 122, 1118–1129.

(7) Devesa SS, Blot WJ, Fraumeni JF Jr. Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. Cancer. 1998; 83, 2049–2053.

(8) Lagergren J. Oesophageal cancer and gastro-oesophageal reflux: what is the relationship? Gut. 2004;53, 1064–1065.

(9) Office for National Statistics (ONS). Cancer survival in England: Patients diagnosed 2005-2009 and followed up to 2010. London: ONS. 2010.

(10) Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-tieulent J, Jemal A. Global Cancer Statistics, 2012. CA a cancer J Clin. 2015; 65(2):87–108.

(11) Hoyo C, Cook MB, Kamangar F, Freedman ND, Whiteman DC, Bernstein L, *et al*. Body mass index in relation to oesophageal and oesophago-gastric junction adenocarcinomas: a pooled analysis from the International BEACON Consortium. Int J Epidemiol. 2012 Dec; 41(6):1706–18.

(12) Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J Clin Oncol. 2006 May 10; 24(14):2137–50.

(13) Mowat C, Williams C, Gillen D, Hossack M, Gilmour D, Carswell A, Wirz, A Preston, T & McColl, K. E. Omeprazole, Helicobacter pylori status, and alterations in the intragastric milieu facilitatingbacterial N-nitrosation. Gastroenterology. 2010; 119, 339–347.

(14)Bodger K, Eastwood PG, Manning SI, *et al*. Dyspepsia workload in urban general practice and implications of the British Society of Gastroenterology Dyspepsia guidelines. Aliment Pharmacol Ther.2000; 14: 413–420.

(15)The NHS Information Centre, 2013. National Oesophago-gastric Cancer Audit 2010. (athttp://www.hqip.org.uk/assets/NCAPOP-Library/NCAPOP-2013-14/NOGCA-2013-INTERACTIVE-PDF27062013.pdf).

(16)Prasad GA, Wu TT, Wigle DA, Buttar NS, Wongkeesong LM, Dunagan KT, Lutzke LS, Borkenhagen LS, Wang KK. Endoscopic and surgical treatment of mucosal (T1a) esophageal adenocarcinoma in Barrett's esophagus. Gastroenterology. 2009;137: 815-2.

(17)Miyamoto A, Kuriyama S, Nishino Y, Tsubono Y, Nakaya N, Ohmori K, *et al*. Lower risk of death from gastric cancer among participants of gastric cancer screening in Japan: A population-based cohort study. Prev Med (Baltim). 2007;44(1):12–9.

(18)Pauling L, Robinson AB, Teranishi R, *et al*. Proc Natl. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. Proc Natl Acad Sci USA 1971; 68:2374–2376

(19)Markar, S. R.; Wiggins, T.; Antonowicz, S.; Chin, S. T.; Romano, A.; Nikolic, K.; Evans, B.; Cunningham, D.; Mughal, M.; Lagergren, J.; Hanna, G. B. Assessment of a noninvasive exhaled breath test for the diagnosis of oesophago-gastric cancer. JAMA Oncol 2018;4(7):970-976.

(20)Amal H, Leja M, Funka K, Skapars R, Sivins A, Ancans G, Liepniece-Karele, Kikuste, Lasina, Haick H. Detection of precancerous gastric lesions and gastric cancer through exhaled breath. Gut. 2016;65(3):400-7

(21)Kumar S, Huang J, Abbassi-Ghadi N, Mackenzie HA, Veselkov KA, Hoare JM, Lovat LB, Spanel P, Smith D, Hanna GB. Mass Spectrometric Analysis of Exhaled Breath for the Identification of Volatile Organic Compound Biomarkers in Esophageal and Gastric Adenocarcinoma. Ann Surg. 2015;262(6):981-90

(22)Kumar S, Huang J, Abbassi-Ghadi N, Spanel P, Smith D, Hanna GB. Selected ion flow tube mass spectrometry analysis of exhaled breath for volatile organic compound profiling of esophago-gastric cancer. Anal Chem. 2013;85(12):6121-8

(23)Kumar S, Huang J, Cushnir JR, Spanel P, Smith D, Hanna GB. Selected ion flow tube-MS analysis of headspace vapor from gastric content for the diagnosis of gastro-esophageal cancer. Anal Chem. 2012;84(21):9550-7

(24)Huang J, Kumar S, Abbassi-Ghadi N, Spanel P, Smith D, Hanna GB. Selected ion flow tube mass spectrometry analysis of volatile metabolites in urine headspace for the profiling of gastro-esophageal cancer. Anal Chem. 2013;85(6):3409-16

(25)Ellis AM, Mayhew CA. Proton transfer reaction mass spectrometry: principles and applications: John Wiley & Sons; 2013.

(26)Millington DS, Kodo N, Norwood DL, Roe CR. Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. J Inherit Metab Dis 1990; 13(3): 321-4.

(27)Ho CS, Lam CWK, Chan MHM, *et al*. Electrospray ionisation mass spectrometry: principles and clinical applications. The Clinical biochemist Reviews 2003; 24(1): 3-12.

(28)Beale DJ, Jones OAH, Karpe AV, *et al*. A Review of Analytical Techniques and Their Application in Disease Diagnosis in Breathomics and Salivaomics Research. International journal of molecular sciences 2016; 18(1): 24.

(29)Gushue JN. Chapter 11 - Principles and Applications of Gas Chromatography Quadrupole Time-of-Flight Mass Spectrometry. In: Ferrer I, Thurman EM, eds. Comprehensive Analytical Chemistry: Elsevier 2013; 255-70.

(30)Marriott P, Shellie R. Principles and applications of comprehensive two-dimensional gas chromatography. Analytical chemistry.2002;21_9-10):573-583

(31)Xue Z, Lixin D, Qi X. Gas Chromatography Mass Spectrometry Coupling Techniques.Plant metabolimics.2014;25-44

(32)Tan LT. Review of GC/MS: A Practical User's Guide, 2nd Edition. Journal of Natural Products 2016; 79(10): 2763.

(33)Sanders JK. GC/MS: A Practical User's Guide (McMaster, Marvin; McMaster, Christopher). Journal of Chemical Education 2000; 77(10): 1282.

(34)Fialkov AB, Steiner U, Lehotay SJ, Amirav A. Sensitivity and noise in GC–MS: Achieving low limits of detection for difficult analytes. International Journal of Mass Spectrometry 2007; 260(1): 31-48.

(35)Steeghs MML, Sikkens C, Crespo E, Cristescu SM, Harren FJM. Development of a proton-transfer reaction ion trap mass spectrometer: Online detection and analysis of volatile organic compounds. International Journal of Mass Spectrometry 2007; 262(1): 16-24.

(36)Graus M, Müller M, Hansel A. High Resolution PTR-TOF: Quantification and Formula Confirmation of VOC in Real Time. Journal of the American Society for Mass Spectrometry 2010; 21(6): 1037-44.

(37)Romano A, Hanna GB. Identification and quantification of VOCs by proton transfer reaction time of flight mass spectrometry: An experimental workflow for the optimization of specificity, sensitivity, and accuracy. Journal of Mass Spectrometry 2018; 53(4): 287-95.

(38)Blake RS, Whyte C, Hughes CO, Ellis AM, Monks PS. Demonstration of Proton-Transfer Reaction Time-of-Flight Mass Spectrometry for Real-Time Analysis of Trace Volatile Organic Compounds. Analytical Chemistry 2004; 76(13): 3841-5.

(39)Sender R, Fuchs S, Milo R. Are we really vastly outnumbered? Revisiting the ration of bacterial to host cells in humans. Cell. 2016;164(3):337-40.

(40)Antonella F, Lorenza DG, Giovanni BD, Elisabetta M, Susanna E. Autism spectrum disorders and the gut microbiota. Nutrients.2019;11(3):521

(41)F. Gich, M. A. Janys, M. König, J. Overmann. Enrichment of previously uncultured bacteria from natural complex communities by adhesion to solid surfaces. Environ. Microbiol. 2012; 14: 2984–2997.

(42)Julian R Marchesi,D H Adams, Francesca Fava, Gerben D A Hermes, Gideon M Hirschfield, Georgina Hold, M. Nabil Quraishi, James Kinross, Hauke Smidt. The gut micriobiota and host health: a new clinical frontier. Gut.2016;65:330-339.

(43)Vakil N, Talley N, van Zanten SV, Flook N, Persson T, Björck E, Lind T, Bolling-Sternevald E. Cost of detecting malignant lesions by endoscopy in 2741 primary care dyspeptic patients without alarm symptoms. Clin Gastroenterol Hepatol. 2009; 7: 756–61.

(44)Thorn R, Reynolds D, GreenmanJ. Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro. Journal of Microbiological Methods. 2011;84(2):258-264.

(45)AllardyceR, Hill A, Murdoch D. The rapid evaluation of bacterial growth and antibiotic susceptibility in blood cultures by selected ion flow tube mass spectrometry. Diagnostic Microbiology and Infectious Disease. 2006;55(4):255-261.

(46)JulákJ, Procházková-FranciscіE, StránskáE, RosováV. Evaluation of exudates by solid phase microextraction–gas chromatography. Journal of Microbiological Methods. 2003;52(1):115-122.

(47)Ethan TH, Hang Lu, Tianming Y, Cindy HN. Microbial ecology along the gastrointestinal tract. 2017;32(4):300-313.

(48)Johan D, Mathilda L *et al*. Molecular characterization of the stomach microbiota in patients with gastric cancer and in controls. J Med Microbiology 2008;58:509–16.

(49)Francisco AJ, Flor VJ, Rafael MG, Alejandra M, Javier T, Stomach microbiota composition varies between patients with non-atrophic gastritis and patients with intestinal type of

gastric cancer. Scientific reports. 2014;10:1038

(50)Chang Soo E, Kim BK, Han DS, *et al*. Differences in gastric mucosal microbiota profiling in patients with chronic gastritis, intestinal metaplasia, and gastric cancer using pyrosequencing methods. Helicobacter 2014;19:407–16.

(51)Yalda K, Yakhya D, Bee Hoon P, Chow Goon NG *et al*. Culturable bacterial microbiota of the stomach of Helicobacter pylori positive and negative gastric disease patients. The scientific world journal. 2014;10.1155/610421

(52)Hyun Jin J, Jaeyeon K, Nayoung K, Ji Hyun P, Ryoung HN *et al*. Analysis of gastric microbiota by pyrosequencing: Minor role of bacteria other than Helicobacter pylori in the gastric carcinogenesis. Helicobacter 2016;21:364-374

(53)Elliott D R, Walker A W, O'Donovan M, Parkhill J, Fitzgerald R C. A non-endoscopic device to sample the oesophageal microbiota: a case-control study. Lancet Gastroenterol Hepatol 2017; 2: 32–42

(54)Hanna,GB.; Boshier, PR.; Markar, SR.; Romano, A. Accuracy and methodologic challenges of volatile organic compound- based exhaled breath tests for cancer diagnosis: a systematic review and meta-anlaysis. JAMA Oncol.2018;16:e182815.

(55)Litvak, D. A.; Hwang, K. O.; Evers, B. M.; Townsend, C. M., Induction of apoptosis in human gastric cancer by sodium butyrate. Jr. Anticancer Res 2000; 20, 779-784.

(56)Okabe, S.; Okamoto, T.; Zhao, C. M.; Chen, D.; Matsui, H. Acetic acid induces cell death: an vitro study using normal rat gastric mucosal cell line and rat and human gastric cancer and mesothelioma cell lines. J Gastroenterol Hepatol 2014; 29 Suppl 4, 65-69.

(57)Haddad, J. J. Current opinion on 3-[2-[(2-tert-butyl-phenylaminooxalyl)-amino]-propionylamino]- 4-oxo-5-(2,3,5,6-tetrafluoro-phenoxy)-pentanoic acid, an investigational drug targeting caspases and caspase-like proteases: the clinical trials in sight and recent anti-inflammatory advances. Recent Pat Inflamm Allergy Drug Discov 2013; 7, 229-258.

(58)Tait, E.; Perry, J. D.; Stanforth, S. P.; Dean, J. R. Idenitifcation of volatile organic compounds produced by bacteria using HS-SPME-GC-MS. J Chromatogr Sci 2014; 52, 363-373.

(59) Haick, H.; Broza, Y. Y.; Mochalski, P.; Ruzsanyi, V.; Amann, Assessment, origin, and implementation of breath volatile cancer markers. A. Chem Soc Rev 2014; 43, 1423-1449.

(60)Zhan X, Duan J, Duan Y. Recent developments of proton-transfer mass spectrometry and its applications in medical research. Mass Spectrom Rev. 2013; 32(2), 143-65

(61)Smith, D.; Španěl, P. Selected ion flow tube mass spectrometry for on-line trace gas analysis. Mass Spectrom Rev. 2005; 24(5), 661−700.

(62)Doran, S.; Romano, A.; Hanna, G. B. Optimisation of sampling parameters for standardised exhaled breath sampling. J Breath Res. 2017;12(1):016007.

(63)Romano, A.; Doran, S.; Belluomo, I.; Hanna, G. B. High-throughput breath volatile organic compound analysis using thermal desorption proton transfer reaction time-of-flight mass spectrometry. Anal Chem 2018; 90, 10204-10210.

(64)Lagg, A., Taucher, J. *et al*. Applications of Proton transfer reactions to gas analysis. Int J. Mass Specrrom and Ion Processes. 1994; 134,55

(65)Szumlas, A. Ray, S. *et al*. Hadamard transform ion mobility spectrometry. Anal Chem.2006; 78(13):4474-81.

(66)Kanu, a., Dwivedi, P. Ion mobility-mass spectrometry. J. Mass Spectrom. 43,1.

(67)Hansel, A.; Jordan, A.; Holzinger, R.; Prazeller, P.; Vogel, W.; Lindinger, W. Proton transfer reaction mass spectrometry: on-line trace gas analysis at the ppb level. International Journal of Mass Spectrometry and Ion Processes 1995; 149–150, 609–619.

(68)Su, T. Parametrization of kinetic energy dependences of ion-polar molecule collision rate constants by trajectory calculations. The Journal of Chemical Physics 1994; 100 (6), 4703–4703.

(69)Cappellin, L.; Karl, T.; Probst, M.; Ismailova, O.; Winkler, P. M.; Soukoulis, C.; Aprea, E.; Märk, T. D.; Gasperi, F.; Biasioli, F. A multipurpose sensor with applications in environmental, agri-food and health science. Environmental Science & Technology 2012; 46 (4), 2283–2290.

(70)Puchalska P, Crawford PA. Multi-dimensional Roles of Ketone Bodies in Fuel Metabolism, Signaling, and Therapeutics. Cell Metab.2017; 25(2):262-84

(71)Currie E, Schulze A, Zechner R, Walther TC, Farese RV, Jr. Cellular fatty acid metabolism and cancer. Cell Metab. 2013;18(2):153-61

(72) Wang, T.; Pysanenko, A.; Dryahina, K.; Španěl, P.; Smith, D. Analysis of breath, exhaled via the mouth and nose, and the air in the oral cavity.*J Breath Res.* 2008; *2*, 037013

(73)Hasim, A.; Ma, H.; Mamtimin, B.; Abudula, A.; Niyaz, M.; Zhang, L.W.; Anwer, J.; Sheyhidin, I. Revealing the metabonomic variation of EC using H-NMR spectroscopy and its association with the clinicopathological characteristics. Mol. Biol. Rep. 2012;39(2), 8955-8964

(74)Bonuccelli G, Tsirigos A, Whitaker-Menezes D, Pavlides, Pestell R, Chiavarina B, Frank P, Howell A, Martinez-Outschoorn U, Sotgia F and Lisanti M. Ketones and lactate “fuel” tumor growth and metastasis:Evidence that epithelial cancer cells use oxidative mitochondrial metabolism. Cell Cycle. 2010 Sep 1;9(17):3506-1

(75)Schug ZT, Peck B, Jones DT, Zhang Q, Grosskurth S, Alam IS, Goodwin LM, Smethurst E, Mason S, Blyth K, McGarry L, James D, Shanks E, Kalna G, Saunders RE, Jiang M, Howell

M, Lassailly F, Thin MZ, Spencer-Dene B, Stamp G, van den Broek NJ, Mackay G, Bulusu V, Kamphorst JJ, Tardito S, Strachan D, Harris AL, Aboagye EO, Critchlow SE, Wakelam MJ, Schulze A, Gottlieb E. Acetyl-coA synthease promotes acetate utilization and maintains cancer cell growth under metabolic stress. Cancer Cell. 2015 Jan 12;27(1):57-71

(76)Guelrud M, Herrera I. Acetic acid improves identification of remnant islands of Barrett's epithelium after endoscopic therapy. Gastrointest Endosc. 1998,47(6):5125

(77)Lambert R, Rey JF, Sankaranarayanan R. Magnification and chromoscopy with the acetic acid test. Endoscopy. 2003;35(5):437-45

(78)Shi X, Zheng C, Li C, Hou K, Wang X, Yang Z, Liu C, Liu Y, Che X, Qu X. 4-Phenybutyric acid promotes gastric cancer cell migration via histone deacetylase inhibition-mediated HER3/HER4 up-regulation. Cell Biol Int. 2018 Jan;42(1):53-62

(79)Miyazaki, Yuji & Kikuchi, Kentaro & González-Alva, Patricia & Inoue, H & Noguchi, Y & Tsuchiya, H & Hayashi, Joichiro & Shin, K & Ochiai, Kuniyasu & Kusama, K. Association of Butyric Acid Produced by Periodontopathic Bacteria with Progression of Oral Cancer. Journal of Cancer Science and Therapy. 2010;2(2):26-32.

(80)Stadler S, Stefanuto PH, Brokl M, Forbes SL, Focant JF. Characterization of volatile organic compounds from human analogue decomposition using thermal desorption coupled to comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry. Anal Chem. 2013 Jan 15;85(2):998-1005

(81)Crotti S, Agnoletto E, Cancemi G, Di Marco V, Traldi P, Pucciarelli S, Nitti D, Agostini M.Altered plasma levels of decanoic acid in colorectal cancer new diagnostic biomarker. Anal Bioanal Chem. 2016 Sep;408(23):6321-8

(82)Kakihara Y, Ichihara K. Studies of phenol formation. I. Method of the determination of phenol and its microbial formation from tyrosine and tyrosine derivatives. Med J Osaka Univ. 1953.

(83)Wu WM, Yang YS, Peng LH. Microbiota in the stomach: new insights. J Dig Dis. 2014;15(2):54–61.

(84)Chin, S. T.; Romano, A.; Doran, S. L. F.; Hanna, G. B. Cross-platform mass spectrometry annotation in breathomics of oesophageal-gastric cancer. *Sci Rep* 2018; *8*, 5139.

(85)Wei, X.; Koo, I.; Kim, S.; Zhang, X. Compound identification in GC-MS by simultaneously evaluating the mass spectrum and retention index. Analyst 2014; 139, 2507-2514.

(86)Amann, A.; Miekisch, W.; Schubert, J.; Buszewski, B.; Ligor, T.; Jezierski, T.; Pleil, J.; Risby, T. Analysis of exhaled breath for disease detection. Annu Rev Anal Chem (Palo Alto Calif) 2014; 7, 455-482.

(87)Kasubuchi M, Hasegawa S, Hiramatsu T, Ichimura A, Kimura I Dietary gut microbial metabolites, short-chain fatty acids, and host metabolic regulation. Nutrients 2015; 7, 2839-2849.

(88)Layden, B. T.; Angueira, A. R.; Brodsky, M.; Durai, V.; Lowe, W. L., Short chain fatty acids and their receptors: new metabolic targets. Jr. Transl Res 2013; 161, 131-140.

(89)Yoshii, Y.; Furukawa, T.; Saga, T.; Fujibayashi, Y. Acetate/acetyl-CoA metabolism associated with cancer fatty acid synthesis: overview and application.Cancer Lett 2015; 356, 211-216.

(90)Migita, T.; Ruiz, S.; Fornari, A.; Fiorentino, M.; Priolo, C.; Zadra, G.; Inazuka, F.; Grisanzio, C.; Palescandolo, E.; Shin, E.; Fiore, C.; Xie, W.; Kung, A. L.; Febbo, P. G.; Subramanian, A.; Mucci, L.; Ma, J.; Signoretti, S.; Stampfer, M.; Hahn, W. C., *et al*. Fatty acid synthase: a metabolic enzyme and candidate oncogene in prostate cancer. J Natl Cancer Inst 2009; 101, 519-532.

(91)Carvalho, M. A.; Zecchin, K. G.; Seguin, F.; Bastos, D. C.; Agostini, M.; Rangel, A. L.; Veiga, S. S.; Raposo, H. F.; Oliveira, H. C.; Loda, M.; Coletta, R. D.; Graner, E. Fatty acid synthase inhibtion with Oristat promotes apoptosis and reduces cell growth and lymph node metastasis in a mouse melanoma model. Int J Cancer 2008; 123, 2557-2565.

(92)Rossi, S.; Graner, E.; Febbo, P.; Weinstein, L.; Bhattacharya, N.; Onody, T.; Bubley, G.; Balk, S.; Loda, M. Fatty acid synthase expression defines distict molecular signature in prostate cancer. Mol Cancer Res 2003; 1, 707-715.

(93)Bik EM, Eckburg PB, Gill SR, *et al*. Molecular analysis of the bacterial microbiota in the human stomach. Proc Natl Acad Sci U S A 2006;103:732–7.

(94)Nardone G, Compare D. The human gastric microbiota: Is it time to rethink the pathogenesis of stomach diseases? United European Gastroenterol J 2015;3:255–60.

(95) Modlin I, Sachs G. Acid related diseases—biology and treatment. Philadelphia:Lippincott Williams & Wilkins; 2004;54(10):1508.

(96) Marshall BJ, Warren JR. Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet 1983;321:1273–5.

(97) Lofgren JL, Whary MT, Ge Z, *et al*. Lack of commensal flora in Helicobacter pylori-infected INS-GAS mice reduces gastritis and delays intraepithelial neoplasia. Gastroenterology 2011;140:210–20.

(98) Lertpiriyapong K, Whary MT, Muthupalani S, *et al*. Gastric colonisation with a restricted commensal microbiota replicates the promotion of neoplastic lesions by diverse intestinal microbiota in the Helicobacter pylori INS-GAS mouse model of gastric carcinogenesis. Gut. 2014;63(1):54-63.

(99) Yu G, Gail MH, Shi J, *et al*. Association between upper digestive tract microbiota and cancer-predisposing states in the esophagus and stomach. Cancer Epidemiol Biomarkers Prev 2014;23:735–41.

(100) Gall A, Fero J, McCoy C, *et al*. Bacterial composition of the human upper gastrointestinal tract microbiome is dynamic and associated with genomic instability in a Barrett's esophagus cohort. PLoS One 2015;10(6):e0129055.

(101) Scheline, R. R. Drug metabolism by intestinal microorganisms. J. Pharm.Sci. 1968;57, 2021–2037.

(102) Jia, W., Li, H., Zhao, L., and Nicholson, J. K. (2008). Gut microbiota: a potential new territory for drug targeting. Nat. Rev. Drug Discov. 2008; 7,123–129: 10.1038/nrd2505.

(103) Modlin I, Sachs G. Acid related diseases—biology and treatment. Philadelphia:Lippincott Williams & Wilkins; 2004;54(10):1508.

(104) Hunt RH, Camilleri M, Crowe SE, *et al*. The stomach in health and disease. Gut 2015;64:1650–68.

(105) Monstein HJ, Tiveljung A, Kraft CH, *et al*. Profiling of bacterial flora in gastric biopsies from patients with Helicobacter pylori-associated gastritis and histologically normal control individuals by temperature gradient gel electrophoresis and 16S rDNA sequence analysis. J Med Microbiol 2000;49:817–22.

(106) Carr FJ, Chill D, Maida N. The lactic acid bacteria: a literature survey. Crit Rev Microbiol 2002;28(4):281–370.

(107) Azcarate-Peril MA, Altermann E, Hoover-Fitzula RL, *et al*. Identification and inactivation of genetic loci involved with Lactobacillus acidophilus acid tolerance. Appl Environ Microbiol 2004;70(9):5315–22.

(108) Moreau MC, Ducluzeau R, Raibaud P. Hydrolysis of urea in the gastrointestinal tract of "monoxenic" rats: effect of immunization with strains of ureolytic bacteria. Infect Immun 1976;13(1):9–15.

(109) Fordtran JS, Walsh JH. Gastric acid secretion rate and buffer content of the stomach after eating. Results in normal subjects and in patients with duodenal ulcer. J Clin Invest 1973;52(3):645–57.

(110) Walsh JH, Richardson CT, Fordtran JS. pH dependence of acid secretion and gastrin release in normal and ulcer subjects. J Clin Invest 1975;55(3):462–8.

(111) Sjo¨ stedt, S., Kager, L., Heimdahl, A. & Nord, C. E. Microbial colonization of tumors in relation to the upper gastrointestinal tract in patients with gastric carcinoma. Ann Surg 1988;207, 341–346.

(112) Biarc, J., Nguyen, I. S., Pini, A., Gosse´ , F., Richert, S., Thierse´ , D., Van Dorsselaer, A., Leize-Wagner, E., Raul, F. & other authors. Carcinogenic properties of proteins with pro-inflammatory activity from Streptococcus infantarius (formerly S. bovis). Carcinogenesis 2004;25,1477–1484.

(113) Afra, K.; Laupland, K.; Leal, J.; Lloyd, T.; Gregson, D. Incidence, risk factors, and outcomes of fusobacterium species bacteremia. BMC Infect. Dis. 2013; 13, 264.

(114) Jing C, Jingjuan Z., *et al*, Chemotherapy alters the phylogenetic molecular ecological 73 networks of intestinal microbial communities. 2019; 10.3389.

(115) Fearon, E.R.; Vogelstein, B. A genetic model for colorectal tumorigenesis. Cell 1990; 61, 759–767.

(116) Sharma BK, Santana IA, Wood EC, *et al*. Intragastric bacterial activity and nitrosation before, during, and after treatment with omeprazole. Br Med J 1984;289:717–9.

(117) Wang L, Zhou J, Xin Y, *et al*. Bacterial overgrowth and diversification of microbiota in gastric cancer. Eur J Gastroenterol Hepatol 2016;28(3):261–6.

(118) Deng K, Lin S, Zhou L, Li Y, Chen M, Wang Y, *et al*. High levels of aromatic amino acids in gastric juice during the early stages of gastric cancer progression. PLoS One. 2012 Jan;7(11):e49434.

(119) Crespo E, de Ronde, H, Kuijper, S. *et al*. Potential biomarkers for identification of mycobacterial cultures by proton transfer reaction-mass spectrometry analysis. Rapid Commun. Mass Spectrom. 2012;26 ,679.

(120) Critchley A. D. J, Elliott, T. S, Harrison, G. *et al*. The proton transfer reaction mass spectrometer and its use in medical science: applications to drug assays and the monitoring of bacteria. Int. J. Mass Spectrom . 2004;239,235.

(121) Bunge M, Araghipour, N, Mikoviny, T. On-line monitoring of microbial volatile metabolites by proton transfer reaction-mass spectrometry. Appl. Environ. Microbiol. 2008;74 , 2179.

(122) Bauerfeind P, Garner R, Dunn BE, *et al*. Synthesis and activity of Helicobacter pylori urease and catalase at low pH. Gut 1997;40:25–30.

(123) Nijland R, Burgess, J. G. Bacteria olfaction. Biotechnol. J. 2010;5,974.

(124) Adamsson I, Nord C E, Lundquist P, Sjo stedt, S & Edlund, C. Comparative effects of omeprazole, amoxycillin plus metronidazole on the oral, gastric and intestinal microflora in Helicobacter pylori infected patients. J Antimicrob Chemother. 1999;44, 629–640.

(125) Eckburg P. B, Bik E. M, Bernstein C. N, Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E. & Relman, D. A. Diversity of the human intestinal microbial flora. Science 308, 2005;1635–1638.

(126) Crespo E, Cristescu, S. M, de Ronde H. *et al*. (2011) Proton transfer reaction mass spectrometry detects rapid changes in volatile metabolite emission by Mycobacterium smegmatis after the addition of specific antimicrobial agents. J. Microbiol. Methods.2011;86 , 8.

(127) Goupry, s., Rochut, N., Robins, R. *et al*. Evaluation of solid-phase microextraction for the isotopic analysis of volatile compounds produced during fermentation by lactic acid bacteria. J. Argic. Food Chem. 2000;48, 2222-2227.

(128) Risticevic, S, Lord, H, Gorecki, T, Aruthur, C.L *et al*. Protocol for solid-phase microextraction method development. Nat. Protoc.2010; 122-139.

(129) T. Pluskal, S. Castillo, A. Villar-Briones, M. Orešič. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, BMC Bioinformatics.2010; 11:395.

(130) Chong, J, Soufan, O, Li, C., Caraus, I, Li, S., Bourque, G., Wishart, D.S. and Xia, J. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. Nucleic Acids Research.2018; 46: w486.

(131) Farag, M.A, zhang, H. & Ryu, C.-M. Dynamic chemical communication between plants and bacteria through airborne signals: induced resistance by bacterial volatiles. J. Chem. Ecol.2013; 39, 1007-1018.

(132) Kai, M.J.D., Stanforth, S.P.& Dean, J.R. Identification of volatile organic compounds produced by bacteria using HS-SPME-GC-MS. J Chromatogr. Sci.2014;52, 363-373.

(133) Zhang J, Liu, L., Wei, S., Nagana-Gowda, G.A., Hammoud, Z., Kesler, K.A, Raftery, D. Metabolomics study of oesophageal adenocarcinoma. J Thorac Cardiovasc Surg. 2011; 141, 469-475.

(134) Bone E, Tamm A, Hill M. The production of urinary phenols by gut bacteria and their possible role in the causation of large bowel cancer. Am J Clin Nutr. 1976 Dec; 29(12):1448–54.

(135) Väkeväinen S, Mentula S, Nuutinen H *et al*. Ethanol-derived microbial production of carcinogenic acetaldehyde in achlorhydric atrophic gastritis. Scand J Gastroenterol 2002; 37: 648–55.

(136) Reaves, M.L. Rabinowitz, J.D. Metabolomics in systems microbiology. Cuu. Opin. Biotechnol.2011;22, 17-25.

(137) Farag, M.A, Ryu, C.-M., summer, L.W. & Pare, P.W GC-MS SPME profiling of rhizobacterial volatiles reveals prospective inducers of growth promotion and induced systemic resistance in plants. Phytochemistry 2006;67, 2262-2268.

(138) Kostrouchova, M.,Kostrouch, Z, Kostrouchova, M. Valproic acid, a molecular lead to multiple

regulatory pathways. Folia Biol (Praha) 2007; 53, 37-49.

(139) Hur H, Paik, M. J, Xuan Y, Nguyen D. T, Ham I. H, Yun J, Cho Y. K,,Lee G, Han S. U Quantitaive measurment of organic acids in tissues from gastric cancer patients indicates increased glucose metabolism .PLoS One 2014; 9, e98581.

(140) Matthews G. M, Howarth G. S, Butler R. N Short-chain fatty acid modulation of apoptosis in the Kato III human gastric carcinoma cell line.Cancer Biol Ther 2007; 6, 1051-1057.

(141) Litvak D. A, Hwang K. O, Evers B. M, Townsend, C. M., Induction of apoptosis in human gastric cancer by sodium butyrate. Jr. Anticancer Res 2000;20, 779-784.

(142) Kasubuchi M, Hasegawa S, Hiramatsu T, Ichimura A, Kimura I Dietary gut microbial metabolites, short-chain fatty acids, and host metabolic regulation. Nutrients 2015; 7, 2839-2849.

(143) Ryu, C.M. *et al*. Bacterial volatiles promote growth in Arabidopsis. Proc.Natl.Acad. Sci.2003;100, 4927-4932.

(144) Windey K, De Preter V, Verbeke K, Relevace of protein fermentation to gut health. Mol Nutr Food Res. 2012;56, 184-196.

(145) Smith EA, Macfarlane GT, Enumaration of human colonic bacteria producing phenolic and indolic compounds: effects of pH, carbohydrate availability and retention time on dissimilatory aromatic amino acid metabolism. J appl Bacteriol, 1996; 81(3), 288-302

(146) Smith EA, Macfarlane GT, Formation of phenolic and indolic compounds by anaerobic bacteria in the human large intestine. Microbiology Eco, 1997; 33, 180-188.

(147) Jeremy M. B, John L. T, Gregory J.G, Lubert S, Biochemistry W.H Freeman and Company, New York, 2012.

(148) Klaus M.H, Lisda M.W, The shikimate pathway, Annu. Rev. Plant Physiol. Plant Mol. Biol., 1999; 50, 473–503.

(149) Gelfard D H, Steinberg R A, Escherichia coli mutants deficient in the aspartate and aromatic amino acid minotransferases. J. Bacteriol. 1977; 130 (1), 429-440.

(150) Cotton RG, Gibson F, The biosynthesis of phenylalanaine and tyrosine; enzymes converting chorismic acid into prephenic acid and their relationships to prephenate dehydratase and prephenate dehydrogenase. Biochim. Biophys. Acta, 1965; 100 (1), 76-88.

(151) Vivijs B, Haberbeck LU, Baiye MM, Bernaerts K, Geerard AH, Formate hydrogen lyase mediates stationary-phase deacidifcation and increases survival during sugar fermentation in acetone-producing eneterobacteria. Front Microbiol, 2015; 6, doi: 10.3389/2015.00150.

(152) Nosova T, Jokelainen P, Kaihovaara H, Jousimies S, Sitonen R, Aldehyde dehydrogenase activity and acetate production by aeroboc bacteria representing the normal flora of human large intestine. Alcohol and Alcoholism, 1996; 31 (6), 555–564

(153) Ying L, XI W, Xizhen G, Pingfang T, High production of 3-hydoroxypropionic acid in kelbesiella pneumonia by systematic optimization of glycerol metabolism. Scientific Reports, 2016; 10.1038/srep26932.

(154) Honggen Z, Zhenyu W, Oscar L, Development and validation of a GC-FID method for quantitative analysis of oleic acid and related fatty acids. J Pharm Anal 2015; 5(4), 223-230.

11 SUPPLEMENTARY FILE

GC Parameters				
Column Type	ZB-WAXPlus™			
Column Length	30 metres			
Internal Diameter	0.25 mm			
Film Thickness	0.25 µm			
Carrier Gas	Helium			
Flow Rate	1 mL min^{-1}			
Injection Type	Split			
Split Ratio	20:1			
Injection Volume	1 µL			
Inlet Temperature	250 °C			
Pressure	13.05 psi			
Additional Info	Standards solubilised in 0.5% v/v ACN (aq) + 5 mM PeOH internal std			
Oven	**Rate / °C min^{-1}**	**Temperature / °C**	**Hold Time/ min**	**Run Time/ min**
(Initial)		55	0.00	0.00
Ramp 1	30	170	7.60	11.43
Ramp 2	30	250	0	14.10
Method of Detection	FID			

Analyte	**RT** *(mins)*	**LOD**	**LOQ**
Acetic acid	3.962	0.65 mM	1.98 mM
Butyric acid	4.772	0.34 mM	1.01 mM
Valeric acid	5.415	0.34 mM	1.01 mM
Hexanoic acid	6.237	0.34 mM	1.01 mM
Phenol	7.996	0.08 mM	0.24 – 0.29 mM
m-cresol	9.236	0.08 mM	0.24 – 0.29 mM
p-cresol	9.381	0.08 mM	0.24 – 0.29 mM
m-ethylphenol	11.235	0.08 mM	0.24 – 0.29 mM
p-ethylphenol	11.421	0.08 mM	0.24 – 0.29 mM

Supplementary file 1: Method parameters used for GC-FID analysis.

ICHTB HTA licence: 12275
REC Wales approval: 12/WA/0196

Professor George Hanna
Imperial College London
St Mary's Hospital
Praed Street
Paddington
London
W2 1NY

17th November, 2016

Dear Prof Hanna,

Re: Tissue Bank application number R16035-3A; modified 14/11/2016
Project title: Influence of the microbiome on gastro-oesophageal cancer patients.

I am pleased to confirm that the Tissue Management Committee of the ICHTB has approved the modification of your application for access, changes include amendment
To protocol to include analysis of samples using the PTR-TOF-MS spectrometer.
Project duration remains unchanged.

In order to satisfy HTA tracking requirements, would you please ensure the samples you use are reported to the tissue bank via the online database and sub-collection, along with any publications you produce, that should contain the acknowledgment wording as per the signed MTA.

Yours sincerely,

Professor Gerry Thomas
DI, HTA Research Licence.

Cc: TM/tissue bank file

ICHTB HTA licence: 12275
REC Wales approval: 12/WA/0196

Professor George Hanna
Academic Surgical Department
10th Floor
QEQM Building
St. Mary's Hospital
London
W2 1NY

01/ March/2017

Dear Prof Hanna;

Re: Tissue Bank application number R16035-5A modified on 24/02/2017;
Project title: The influence of the microbiome on Volatile Organic Compound production in gastro-oesophageal adenocarcinoma

I am pleased to confirm that the Tissue Management Committee of the ICHTB has approved your amended application for access, modification for the *further analysis*
60 samples (cancer and non-cancer samples) duration of project remains unchanged.

In order to satisfy HTA tracking requirements, would you please ensure the samples you use are reported to the tissue bank via the online database and sub-collection, along with any publications you produce, that should contain the acknowledgment wording as per the signed MTA.

Yours sincerely,

Professor Gerry Thomas
DI, HTA Research Licence.

Cc: TM/tissue bank file

Imperial College Healthcare Tissue Bank
Department of Surgery and Cancer
Charing Cross Hospital,
Fulham Palace Road,
London W6 8RF
Tel: 0203 311 7173
Email: tissuebank@imperial.ac.uk

ICHTB HTA licence: 12275
REC Wales approval: 12/WA/0196

Professor George Hanna
Academic Surgical Department
10th Floor
QEQM Building
St. Mary's Hospital
London
W2 1NY

15th May 2017

Dear Professor Hanna

Re: Tissue Bank application number; R16035-6A

Project title; The influence of the microbiome on Volatile Organic Compound production in gastro-oesophageal adenocarcinoma

I am pleased to confirm that the Imperial College Healthcare Tissue Bank has received and approved the below mentioned amendment to your project.

- An increase in the number of samples requested to 200 samples (cancer and non-cancer samples) in total, each of which will be used to measure tissue VOCs using the instrument in the application.
- Extension of further 36 months (new expiry date; 14th May 2020).

In order to satisfy HTA tracking requirements, would you please ensure the samples you use are reported to the tissue bank via the online database and sub-collection, along with any publications you produce, that should contain the acknowledgment wording as per the signed MTA.

Yours sincerely,

Professor Gerry Thomas
DI, HTA Research Licence.

Imperial College London